PNNL-14931

Pacific Northwest National Laboratory
Operated by Battelle for the U.S. Department of Energy

Statistical Methods and Tools for UXO Characterization

SERDP Final Technical Report

SERDP Project Number: UXO 1199

Lead Principal Investigator: Brent A. Pulsipher

B. A. Pulsipher
R. O. Gilbert
J. E. Wilson
N. L. Hassig
R. F. O'Brien
D. J. Bates
G. A. Sandness
K. K. Anderson

November, 2004

Revision #0

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
NOV 2004		**00-00-2004 to 00-00-2004**

4. TITLE AND SUBTITLE
Statistical Methods and Tools for UXO Characterization

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Pacific Northwest National Laboratory,Richland,WA,99352

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**67**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062; ph: (865) 576-8401 fax: (865) 576-5728 email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161 ph: (800) 553-6847 fax: (703) 605-6900
email: orders@ntis.fedworld.gov online ordering: http://www.ntis.gov/ordering.htm

This document was printed on recycled paper.
(9/2003)

PNNL-14931

Statistical Methods and Tools for UXO Characterization

SERDP Final Technical Report

SERDP Project Number: UXO 1199

Lead Principal Investigator: Brent A. Pulsipher

B. A. Pulsipher
R. O. Gilbert
J. E. Wilson
N. L. Hassig
R. F. O'Brien
D. J. Bates
G. A. Sandness
K. K. Anderson

November, 2004

Revision #0

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

Contents

Figures

Tables

Acronyms

BBR	Badlands Bombing Range
BRAC	Base Realignment and Closure
CSM	Conceptual Site Model
DoD	U.S. Department of Defense
DOE	U.S. Department of Energy
DHS	U.S. Department of Homeland Security
DQO	Data Quality Objectives
EPA	U.S. Environmental Protection Agency
ESTCP	Environmental Security Technology Certification Program
FAR	False Alarm Ratio
FUDS	Formerly Used Defense Site
OE	Ordnance and Explosives
ORNL	Oak Ridge National Laboratory
MLE	Maximum Likelihood Estimate
MTS	Mitretek Systems, Inc.
PNNL	Pacific Northwest National Laboratory
TA	Target Area
TDMD	Time Domain Metal Detector
UXO	Unexploded Ordnance
VSP	Visual Sample Plan Software

Acknowledgements

It is a pleasure to acknowledge Dr. Anne M. Andrews, Program Manager for UXO for the Environmental Security Technology Certification Program (ESTCP) and the Strategic Environmental Research and Development Program (SERDP) for her leadership and enthusiastic support of this PNNL project. Within PNNL the following statisticians and scientists have made valuable technical contributions: Dick Gilbert, Kevin Anderson, Rick Bates, Deborah Carlson, Nancy Hassig, Pat Heasler, Craig McKinstry, Bob O'Brien and Gerald Sandness. Special appreciation is also extended to Luci Walker and Nell Cliff of PNNL for their steadfast and efficient administrative assistance, including financial accounting and the preparation of SERDP conference posters and project reports.

There are several organizations and individuals not affiliated with PNNL that provided valuable contributions to our work. We are particularly happy to acknowledge the close collaborative working relationship we have enjoyed with Sean McKenna of the Sandia National Laboratory (SNL). Our mutual discussions on technical aspects of UXO density estimation methods and geostatistics have been very beneficial. This relationship has evolved to the point that PNNL and SNL are now a team working together on a funded ESTCP project to demonstrate at real DoD training facilities an integrated PNNL/SNL approach to UXO characterization.

PNNL appreciates very much the work that Mike Simmons of Mitretek, Inc. has conducted to generate and provide us with simulated anomaly data at hypothetical sites so that we could determine the performance of our target area detection algorithms. He has been very helpful and responsive to our needs. We would also like to thank Bill Doll, David Bell and George Ostrouchov of the Oak Ridge National Laboratory (ORNL) for the efforts they made to discuss with us their knowledge and experience of working on UXO site characterization issues. Also, we would like to acknowledge the assistance provided by Laura Wrench of Versar, Inc. Laura helped us to "come up to speed" on UXO terminology, concepts, and site characterization methods and how the conceptual site model and systematic planning process were being used or planned for UXO work. She has been a valuable resource and peer reviewer of our ideas and work throughout the project. We also acknowledge the important contributions of all members of the Technical Advisory Committee (TAC) for their insightful peer reviews of the statistical methods developed by this project. Their comments have guided us in many ways.

Executive Summary

The Strategic Environmental Research and Development Program (SERDP) issued a statement of need for FY01 titled *Statistical Sampling for Unexploded Ordnance (UXO) Site Characterization* that solicited proposals to develop statistically valid sampling protocols for cost-effective, practical, and reliable investigation of sites contaminated with UXO; protocols that could be validated through subsequent field demonstrations. The SERDP goal was the development of a sampling strategy for which a fraction of the site is initially surveyed by geophysical detectors to confidently identify clean areas and subsections (target areas, TAs) that had elevated densities of anomalous geophysical detector readings that could indicate the presence of UXO. More detailed surveys could then be conducted to search the identified TAs for UXO.

SERDP funded three projects: those proposed by the Pacific Northwest National Laboratory (PNNL) (SERDP Project # UXO 1199), Sandia National Laboratory (SNL), and Oak Ridge National Laboratory (ORNL). The projects were closely coordinated to minimize duplication of effort and facilitate use of shared algorithms where feasible. This final report for PNNL Project 1199 describes the methods developed by PNNL to address SERDP's statement-of-need for the development of statistically-based geophysical survey methods for sites where 100% surveys are unattainable or cost prohibitive.

This PNNL project focused on 4 objectives:

- Finding target areas of anomalies
- Compliance assessment
- Dig list stopping rules
- Validation of statistical tools developed for TA detection

Figure 1 shows the relationship between these objectives and provides a brief summary of the main results, including developed software and sources of additional information.

Objective 1: Finding Target Areas of Anomalies

This project developed statistically based designs for geophysical detector surveys conducted along transects (swaths) to achieve a specified high probability of traversing and detecting target areas (TAs) of critical size, shape, and density of detectable items of interest (anomalies), some of which may be UXO. These methods were developed for when the density of anomalies in the TA has either a Uniform distribution (unchanging density over space) or a Bivariate Normal distribution for which the anomaly density increases in the shape of a two-dimensional bell curve to a location of maximum density.

Our sampling design methods are consistent with systematic planning processes such as the Data Quality Objectives (DQO) process. The target detection design methods focus on determining the number and spacing of transects (swaths) that will be required to have an acceptably high probability of traversing and detecting target areas of a critical size, shape, and anomaly density. Target areas are defined as areas where UXO are most probable and the number of geophysical anomalies is much higher than in the background.

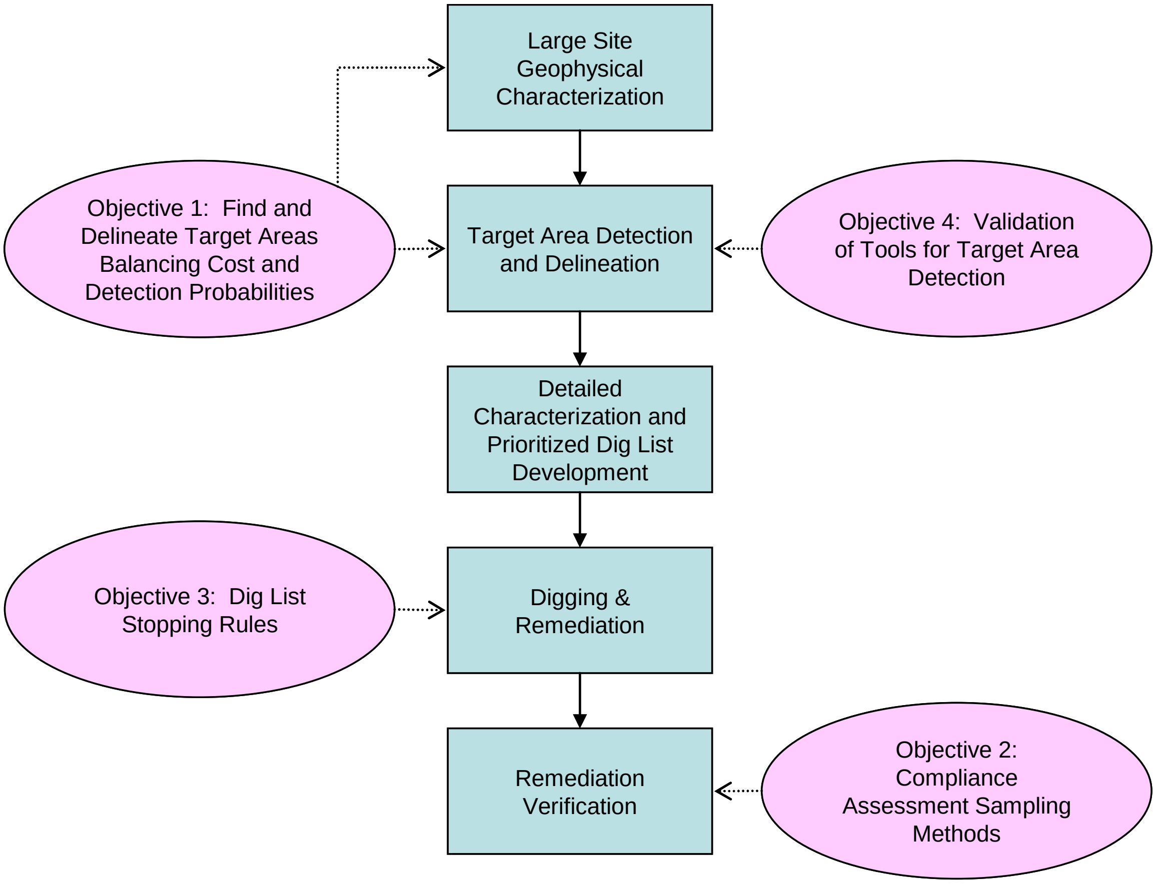

Figure 1. Objectives of the PNNL Project "Statistical Methods and Tools for UXO Characterization"

There are several design parameters (DQOs) that affect the probability of detecting a target zone of interest and the feasibility of the sampling design. These include:

- The assumed shape (circular or elliptical) and size (area in m^2 or ft^2) of a target area that must be detected with high probability.
- The required probability that at least one transect traverses some portion of the assumed target area.
- The swath width or footprint of the geophysical sensor used to detect anomalies.
- The transect pattern (we assume a parallel, square grid, or rectangular grid pattern).
- The spatial orientation (N-S, E-W) of the transect pattern with respect to expected orientation of the target area.
- The anomaly density patterns [we currently assume either a uniform pattern (unchanging density over the target area) or a bivariate normal pattern (highest density at the center of the target area with decreasing density in all directions)].
- The trigger anomaly density threshold value above which we will flag an area as being a potential target area.
- The critical anomaly density threshold value above which we must have a very high probability of flagging as a target area.
- The false negative detection error rate of the geophysical sensor instrument to be used.

- The costs for transect and geophysical sensor setup and sampling/evaluation costs per linear foot of transect.

The effect of varying each of these parameters can be evaluated and the most feasible, cost-effective sampling design can be determined.

Given the design parameters, VSP will compute the required spacing of transects to achieve the specified probability of traversing the critical TA. VSP also calculates the probability of both traversing and detecting the critical TA if it exists, lays the transects down on the site map and outputs the x, y coordinates of the transects. The user can then conduct a sensitivity analysis by evaluating the effects of the input parameters and their required DQOs on VSP outputs. These methods and tools allow the project team to balance DQOs against costs and other site constraints. After analysis of the transect sampling data, if potential TAs are identified, then the next phase of sampling may focus on defining the boundary of the TA and developing a sufficiently accurate map of the anomaly or UXO density. If no TAs are identified, the actions agreed upon and documented in site planning documents, e.g., "no further action" or long term monitoring, may be implemented.
This effort included also developing

- post-survey TA detection evaluation methods
  - to approximate the probability of both traversing and detecting a TA with meandering transects when the density of anomalies in the TA follows a Uniform or Bivariate Normal distribution
  - to analyze anomalies detected along transects to determine the boundary (perimeter) of areas within which TAs are most likely to exist
- a Bayesian method to compute the probability that a TA exists when none was found using a geophysical survey

These methods are described in Gilbert et al., (2003a) and have been coded into the Visual Sample Plan (VSP) software. VSP is being developed by PNNL with funding from the U.S. Department of Energy (DOE), the U.S. Environmental Protection Agency (EPA), the U.S. Defense Department (DoD), and the U.S. Department of Homeland Security (DHS). VSP and its User Guide are available for free download from http://dqp.pnl.gov/vsp.

The demonstration and validation of these methods is discussed under Objective 4 below.

Objective 2: Compliance Assessment

This objective focused on developing statistical methods for determining the number and location of transects for conducting geophysical surveys to achieve high confidence that few if any UXO are present in un-surveyed portions of sites that are (1) very large and not expected to contain UXO or (2) at sites or portions of sites that have undergone remediation to remove UXO. Two statistical design methods from the statistical literature were adapted to achieve this objective:

- Schilling's method (1982), which is used to determine the fraction of an area that should be surveyed and be found to contain no UXO in order to be sufficiently confident that no more than an acceptably small fraction of the entire area could contain UXO, and
- the Wright/Grieve method (Wright 1992, Grieve 1994) to determine the fraction of an area that should be surveyed and be found to contain no UXO in order to be sufficiently confident that the entire area contains *no* UXO.

The desired confidence statements for both methods can only be made if the geophysical survey finds no UXO in the transects that are surveyed. Schilling's method is an acceptance (compliance) sampling plan that is commonly used in industry to assess whether batches (lots) of manufactured products have a defect rate that is acceptably low. Acceptance sampling methods are currently in use by DoD for other purposes. For example, Part B of DoD (1999) provides guidance on the use of acceptance (compliance) sampling for product acceptance in support of the military standard MIL-STD-1916.

The Wright/Grieve method is a statistical Bayesian approach originally developed for where it is important to be able to declare with high probability that there are *no* "defectives" present. For UXO application, a given survey transect (swath) is "defective" if it contains UXO or indications that UXO may be present. Both the Schilling and Wright/Grieve methods are coded into the VSP software for the case where each survey transect over the area of interest has approximately the same width and length.

The Schilling and Wright/Grieve methods have not been validated by application at actual training ranges or using simulated data. Although the methods look promising and are currently available for use in the VSP software tool, the methods should be validated by demonstrating their effectiveness at actual sites. An assumption of both methods is that the geophysical sensor has negligibly small false negative and false positive error rates (probabilities). Of course, this assumption is often not realistic in practice. One way of mitigating the effects of false negative sensor error rates on the Schilling and Wright/Grieve methods is to compensate by surveying a larger fraction of the area than would otherwise be indicated by these methods. A possible method of doing this for Schilling's method is given in Section 4.2.1 of this report. Additional attention to developing methods to adjust or compensate for false negative and false positive sensor error rates is recommended.

Objective 3: Dig List Stopping Rules

The purpose of Objective 3 was to develop a statistical Bayesian methodology for deciding if and when to stop digging locations on the anomaly "dig list" such that there is only a small probability that un-dug locations on the list contain UXO. Using the Bayesian methodology, if no UXO are found at the dug locations, then one can state with confidence that there is very little chance that remaining un-dug locations on the dig list contain UXO. PNNL developed a promising maximum likelihood and Bayesian method, which is described in the "white paper" requested by SERDP (O'Brien, Heasler, and Anderson 2003). The proposed method is also summarized later in this report (Section 4.3). Further work to assess the practicality of the methodology over a broad range of sites and incorporate the methods into VSP has been funded by ESTCP.

Objective 4: Validation of Statistical Methods and Tools

The purpose of Objective 4 is to demonstrate the applicability and validity of the TA detection tools developed by PNNL for Objective 1. This was accomplished by working closely with Mitretek Systems, Inc. (MTS). SERDP funded MTS to develop the SimRangE computer code for generating virtual range environments (spatial patterns of anomalies detected by geophysical surveys) that are representative of typical military ranges (Simmons 2003).

The evaluations of PNNL's tools were conducted during three validation exercises over two phases. These evaluations were preceded by a data exchange phase to validate the format of the

data exchanged between MTS and PNNL and to identify the level of detail needed by PNNL for input into its statistical tools.

Phase I built upon the basic SimRangE model that was used in the data exchange trial. The first validation exercise in Phase I incorporated features such as background clutter along with 155mm projectiles and 4.2 inch mortar weapon systems. The second validation exercise included modeling other weapons systems and delivery methods in order to evaluate the ability of the PNNL tools to differentiate between TAs with different dispersion characteristics.

Phase 2 models incorporated and considered the effects of varying terrain, soil conditions, and topology and their effect on munition dispersion and sensor performance. The features added additional complexity to the overall model and represented a more realistic representation of the challenges that might be encountered in a field sampling study conducted at an actual military range. Phase 2 also included conducting a third validation exercise. This exercise evaluated the ability of PNNL's tools to delineate between frag and clutter of varying densities.

The PNNL tools performed very well on these three validation exercises. The tools identified 100% of the TAs. The tools also delineated boundaries that enclosed 100% of the UXO generated by SimRangE. The PNNL tools also found isolated TAs and were shown to be effective at handling multiple transect survey areas and detection criteria.

It was noted previously that SERDP funded the SNL to develop statistical tools. The SNL tools were also validated using the simulated data generated by the SimRangE model. A demonstration plan for the fourth and final evaluation of statistical tools using the SimRangeE model demonstration plan has been prepared for approval by the ESTCP. This demonstration and evaluation will apply to a proposed integrated PNNL/SNL approach wherein the methods and tools of both PNNL and SNL are combined into a logical framework for improved and more complete characterization of anomaly density and the detection of UXO TAs.

1.0 Objectives

The SERDP statement of need for FY01 titled "Statistical Sampling for Unexploded Ordnance (UXO) Site Characterization" solicited proposals to develop statistically valid sampling protocols for cost-effective, practical, and reliable investigation of sites contaminated with UXO; protocols that could be validated through subsequent field demonstrations. The SERDP goal was the development of a sampling strategy for which a fraction of the site is initially investigated (sampled) to identify clean and contaminated subsections (TAs) with high confidence. This strategy would appropriately direct detailed geophysical surveys in search of the UXO TAs. The sampling protocols developed were to consider appropriate geophysical instrumentation, including towed arrays, which can cost-effectively sample extended linear transects (swaths). Also, the methods developed were to apply to a variety of site conditions with regard to geology, climate, terrain, vegetation, and history of use, and should consider various potential ordnance contamination (bombing targets to small arms ranges).

SERDP funded the projects proposed by the Pacific Northwest National Laboratory (PNNL), Sandia National Laboratory (SNL) and Oak Ridge National Laboratory (ORNL). The projects were closely coordinated to minimize duplication of effort and facilitate use of shared algorithms where feasible. This final report focuses on the PNNL project.

In response to the SERDP statement of need, the technical objective of this PNNL project was to develop statistically defensible sampling strategies and data analysis methods and tools to

- guide planning of UXO site characterization and clean-up verification efforts, and to
- permit balancing the risk and confidence in not finding existing UXO TAs versus unnecessary or non-optimal searches for UXO.

The initial objectives (deliverables) of the PNNL project were:

Finding Target Areas of Anomalies

Develop a sampling strategy wherein straight-line or meandering transects (swaths) are surveyed by geophysical detectors to achieve a high probability of detecting TAs of critical size, shape, and density of detectable items of interest (anomalies), some of which may be UXO.

Compliance Assessment

Develop a strategy for conducting geophysical surveys along transects or in geographical units (areas) to achieve high confidence that no or very few UXO are present at large sites believed to have low potential for UXO, or that UXO remain at sites from which UXO have been removed during UXO cleanup operations.

In FY03 the following two objectives were added to the PNNL project:

Dig List Stopping Rules

Develop a sampling strategy, data analysis methods, and a stopping rule for when to stop digging locations on the "dig list" while achieving confidence that un-dug locations on the list do not contain UXO. If no UXO are found at the dug locations, the developed methods should allow one to state with confidence that there is very little chance that remaining un-dug locations on the dig list contain UXO.

Validation of Statistical Methods and Tools

Use a simulated range environmental model (SimRangE) developed by Mitretek Systems, Inc. (MTS) to exercise and validate the statistical tools developed by PNNL for detecting and delineating UXO TAs.

2.0 Background

Past military training and weapons testing activities have resulted in UXO being present at sites designated for base realignment and closure (BRAC) and at Formerly Used Defense Sites (FUDS). It is not practical or economically feasible to conduct 100% geophysical surveys over the millions of acres of land that are potentially contaminated with UXO.

Defensible statistical sampling strategies for surveying less than 100% of these areas are needed. In response, statistical software tools collectively known as *SiteStats/GridStats* were developed for UXO sampling and characterization by the U.S. Army Corps of Engineers for the U.S. Army. The software sought to guide sampling of underground anomalies to estimate the density of UXO and to delineate areas of homogeneous UXO density that can be used in the formulation of response actions. However, problems arose in the use of *SiteStats/GridStats*. A technical evaluation of the software tools by Ostrouchov, G., et al., (1999) states:

> "It was found that *SiteStats/GridStats* does adequately guide the sampling so that the UXO density estimator for a sector is unbiased. However, the software's techniques for delineation of homogeneous areas perform less well than visual inspection, which is frequently used to override the software in the overall sectorization methodology. The main problems with the software lie in the criteria used to detect nonhomogeneity and those used to recommend the number of homogeneous subareas."

SiteStats/GridStats was also evaluated by Singh, Engelhardt and Singh (2000, 2001). They found that *SiteStats/GridStats* yielded incorrect estimates of UXO densities and that one of the main reasons for this failure was the unrealistic assumption of homogeneous and uniform anomaly or UXO distribution within the area under study.

These problems motivated the work of this PNNL project.

3.0 Methods and Materials

PNNL organized a team of statisticians and other support personnel to achieve the objectives of this PNNL project. The goal for the team was to develop methods that could be applied at specific DoD sites in the context of an iterative systematic planning process such as the Data Quality Objectives (DQO) process (USEPA 2000) and to code the developed methods into the VSP software tool. Past and current sponsors of VSP are the U.S. Department of Energy (DOE), the U.S. Environmental Protection Agency (EPA), the U.S. Department of Defense (DoD), and the U.S. Department of Homeland Security (DHS). The VSP software and user's guide (Hassig et al., 2002) are available for free download at http://dqo.pnl.gov/vsp.

Systematic planning for a site survey begins by carefully defining the goals and objectives of the geophysical survey, constructing the conceptual site model (CSM) (Lantzer et al., 2001), which is updated as additional information is obtained over time, and specifying the DQOs. The CSM is particularly important because it provides the basis and rationale for the site-specific survey transect design developed via the systematic planning process.

3.1 Objective 1: Finding Target Areas

Initial activities included obtaining maps and other information that may be available about a site from an archive search report. Information was obtained about the likely shape of TAs, firing points, and the safety fans, as well as CSM components that affect the UXO survey design. This and other information obtained from persons knowledgeable about UXO characterization efforts at DoD training sites were used to learn the critical issues and practical constraints about geophysical surveys for anomalies and UXO.

Next, we developed the mathematical methods needed to determine the amount of geophysical surveying required at a site to traverse and detect TAs of anomalies with sufficient confidence, as described in Gilbert et al. (2003a). First, equations were developed to compute the maximum spacing between transects to achieve a required probability that a TA of specified size, shape, and orientation is traversed by one or more transects of specified width. Equations were developed for three different transect patterns: parallel, square grid, or rectangular grid for circular or elliptical TAs. Then, methods were developed for computing the probability that the geophysical detector not only traverses but also *detects* the TA, as documented in Section 3.0 of Gilbert et al. (2003a).

Methods were developed for two different assumptions about the model of the density of anomalies over the spatial extent of the TA: a Uniform distribution (unchanging density throughout the TA) and a Bivariate Normal distribution (density increases in the shape of a two-dimensional bell curve from the edge to the center of the TA). The probability that the geophysical sensor detects the TA given that the TA was traversed was modeled as a binomial distribution where the probability that an anomaly in a transect being surveyed is not detected by the geophysical detector is the false negative detection error rate of the detector, which is assumed known. This rate is assumed constant for all anomalies that are traversed. A computer simulation algorithm was developed so that the randomness in the placement of the grid of transects over the site was taken into account. Different placements will change the extent that the TA is traversed by one or more transects, which will change the number of anomalies that will be scanned by the geophysical detector, which will change the probability that the TA will be detected. All of these methods were coded into the VSP software so that the VSP user can easily

determine the spacing between transects required to achieve the desired probability of traversing and detecting the TA. In addition, a Bayesian method for computing the probability that a TA exists when none has been detected was also developed and coded into VSP (Gilbert et al, 2003a, Section 4.0), based on methodology in Gilbert (1987, pages 128-129).

Next, methods for detecting TAs were developed to take into account that field crews may find that obstacles, dense vegetation, geographic features, and other factors can make it undesirable, difficult, or impossible to follow straight-line transects with geophysical sensors. The methods, which were coded into VSP, approximate the probability that a TA would have been found using the straight-line or meandering transects that were *actually used* in the geophysical survey (Gilbert et al. 2003a, Sections 5.1, 5.2 and 5.3) when the density distribution within the TA is Uniform or Bivariate Normal.

The methods developed for this objective were validated as described in Section 3.4 of this report.

3.2 Objective 2: Compliance Assessment

The statistical design methods by Schilling (1982), Wright (1992) and Grieve (1994) were evaluated and selected. These methods are proposed here for DoD sites where UXO remediation efforts have occurred or at very large DoD sites or portions of sites that are not expected to have UXO. At the request of SERDP, a compliance assessment white paper (Gilbert et al., 2003b) was written that describes the methods and how they can be used, as discussed later in this report (Section 4.2). The design methods determine the proportion of the site that must be surveyed and found to contain no UXO in order to achieve a required high confidence that few if any UXO are present in the un-surveyed area. A proposed modification to Schilling's method to account for false negative sensor error rates was developed.

3.3 Objective 3: Dig List Stopping Rules

A Bayesian methodology based on the Bayes Rule (Casella and Berger 1990) was proposed. In practice, based on the analysis of anomalous readings obtained by a geophysical survey of the site, a dig list of N locations of anomalous readings is created. These locations are assigned to k bins ranging from "most likely to contain UXO" to "least likely to be UXO" based on signal discrimination techniques and expert judgment. In the absence of a stopping rule, all locations on the dig list are dug to determine if UXO is present at each location. While this method is thorough and provides near 100% certainty that all locations on the dig list that contain UXO are identified, it can also be very expensive. PNNL developed a statistical Bayesian methodology that can result in digging less than 100% of locations on the dig list while achieving a pre-specified low probability that any of the un-dug locations on the dig list contain UXO. At the request of SERDP, a "white paper" on the proposed dig list methodology (O'Brien, Heasler and Anderson 2003) was written. This methodology is described in detail in Section 4.3 of this report, and more briefly in the next section.

3.4 Objective 4: Validation of Statistical Tools

Mitretek Systems, Inc. (MTS) developed a SimRangE Development and Analysis Plan (Simmons 2003) to exercise and validate the statistical tools developed by PNNL. The goal of the SimRangeE exercise was to develop virtual range environments representative of typical military

ranges. SimRangeE was used to develop, analyze, and validate the statistical methods and tools developed by this PNNL project for enhancing UXO site characterization.

SimRangE is a numerically based model that was developed primarily using Microsoft© Excel. The evaluations of PNNL's tools were conducted in two phases, which were preceded by a data exchange phase to validate the format of the data exchanged between MTS and PNNL and to identify the level of detail needed for input into PNNL's statistical tools that had been developed.

Phase I built upon the basic SimRangE model that was used in the data exchange trial. The first exercise in Phase I incorporated features such as background clutter along with 155 mm projectiles and 4.2 inch mortar weapon systems. The second exercise included modeling other weapons systems and delivery methods in order to evaluate the ability of the PNNL methods and tools to differentiate between TAs with different dispersion characteristics.

Phase 2 models considered the effects of varying terrain, soil conditions, and topology and their effect on munition dispersion and geophysical sensor performance. The features added additional complexity to the site model and represented a more realistic representation of the challenges that might be encountered in surveying an actual military range with geophysical sensors. Two exercises were conducted to evaluate the ability of PNNL's methods and tools to delineate between frag and clutter of varying densities. Additional information on the goals and design of SimRangE is provided by Simmons (2003)

Each validation exercise began by MTS constructing a model of the effects on a simulated training range of conducting military training activities using various weapons systems over a long period of time. MTS supplied PNNL with a site map and some information relative to past site practices. Based on required detection probability rates and likely fragmentation dispersion patterns, PNNL determined the optimal transect spacing and selected a set of geophysical transects. MTS used the SimRangE software to generate all locations along the requested transects where anomalies were detected. PNNL applied the VSP density search algorithm to these anomalies to identify the location of TAs.

The VSP search algorithm searches along all the transects checking each search window for an elevated anomaly density. If the density exceeds the trigger density specified by the VSP user, then the center of the window is marked on the site map. The VSP user can also specify a "change in density search." For this option, if the change in density of adjacent windows exceeds the trigger density, then the center of the window is marked on the map. When the search is complete, a histogram of the frequency of windows at various densities is displayed.

The VSP user can specify either a circular or linear search window. A circular window radiates out from the window center and may include more than one transect. The linear search window extends down a single swath and is the width of the swath itself. The VSP user inputs the length of the linear window or the diameter of the circular window. The algorithm steps 1/6 this distance for each new window to check for an elevated anomaly density. The VSP code will plot on the site map all search windows for which the density exceeds the trigger density. Alternatively, the VSP user can specify that VSP color each window based on the density present in the window.

4.0 Results and Accomplishments

4.1 Objective 1: Finding Target Areas of Anomalies

PNNL's methods for detecting TAs of anomalies allows stakeholders to explicitly balance the risk of not finding areas of potential UXO contamination when they may be present and the cost of unnecessary or non-optimal searches. It is assumed that geophysical sensors will be used along transects laid over the site to look for TAs, i.e., areas where the anomaly density is greater than the background density; an indication that UXO may be present. Objective 1 focuses on designing a transect scheme that will have a high probability of traversing and detecting a higher than normal anomaly density area that would be characteristic of the size, shape, and density that might be expected for a TA of concern. At the design stage it is not known whether the anomalies represent clutter, scrap, ordnance and explosives (OE), or UXO. Detected TAs would require further characterization and digging to identify any UXO that are present.

Separate surveys, perhaps using different geophysical sensors or distances between transects, may be conducted for subsections of the site. For example, based on the type of activities that occurred in a subsection of a military base, it may be expected to contain no TAs, whereas another subsection may be expected or known to contain one or more TAs, perhaps at unknown locations. The amount of geophysical surveying needed will typically vary from section to section. The division of a site into subsections should be made on the basis of expert opinion and the CSM. The design methods developed in this PNNL project can then be separately applied to each subsection.

4.1.1 Straight-Line Transects

PNNL first considered geophysical surveys that would be conducted along straight line transects (swaths) laid out in parallel or in a square or rectangular grid pattern. Later, the use of meandering transects was considered, as discussed in Section 4.1.2.

The TA detection design methods focus on determining the number and spacing of transects required to achieve an acceptably high probability of traversing and detecting a TA of anomalies of a critical size, shape, and anomaly density. The methods developed, which are used in the VSP software tool discussed later in this section, are documented in Gilbert et al. (2003a). Section 2.0 and Appendix A of that report provide the equations developed to compute the spacing between transects that is required to traverse circular or elliptical TAs with specified high probability. Section 3.0 of Gilbert et al. (2003a) provides the following equation developed by PNNL to compute the probability, P_{TD}, of traversing and detecting a TA:

$$P_{TD} = (1 - \beta)\overline{P}_{D|T} \quad (1)$$

where

$1 - \beta$ = probability required that the TA is traversed by one or more transects

$\overline{P}_{D|T}$ = mean probability that the survey detects the TA given that the TA is traversed by one or more transects

The algorithms developed to compute $\overline{P}_{D|T}$ are provided in Gilbert et al. (2003a, pages 8-9). These algorithms depend on several design parameters (DQOs):

- The assumed shape (circular or elliptical) and size (area in m^2 or ft^2) of a TA that must be detected with high probability.
- The density of anomalies (number/unit area) in the TA.
    - Two anomaly density distributions were considered: the Uniform and Bivariate Normal distributions.
    - The Uniform distribution applies when the anomaly density remains the same throughout the TA.
    - The Bivariate Normal distribution applies when the anomaly density is lowest at the edge of the TA and increases to a maximum at the center of the TA. An example of a bivariate normal distribution is shown in Figure 2.
- The probability required that at least one transect traverses the TA to some extent.
- The transect width ("footprint" of the geophysical sensor used).
- The transect pattern that is specified (parallel, square grid, or rectangular grid pattern).
- The spatial orientation (N-S, E-W) of the transect pattern with respect to the expected orientation of the TA.
- The site-specific *trigger* anomaly density value above which an area is flagged as being a potential UXO TA. The trigger density is the lowest anomaly density of concern, that is, there is no need to find a TA that has an anomaly density that is less than the trigger density. In practice, the trigger density might be equal to or somewhat larger than the background anomaly density of clutter and non-OE materials.
- The site-specific *critical* density value, which is the anomaly density for which an area must be flagged as a potential UXO TA with a very high probability so that the TA can be searched more closely for UXO. The critical density is larger than the trigger density.
- The false negative detection error rate (probability) of the geophysical sensor that will be used at the site.
- The costs per linear foot of transect for transect and geophysical sensor setup and surveying.

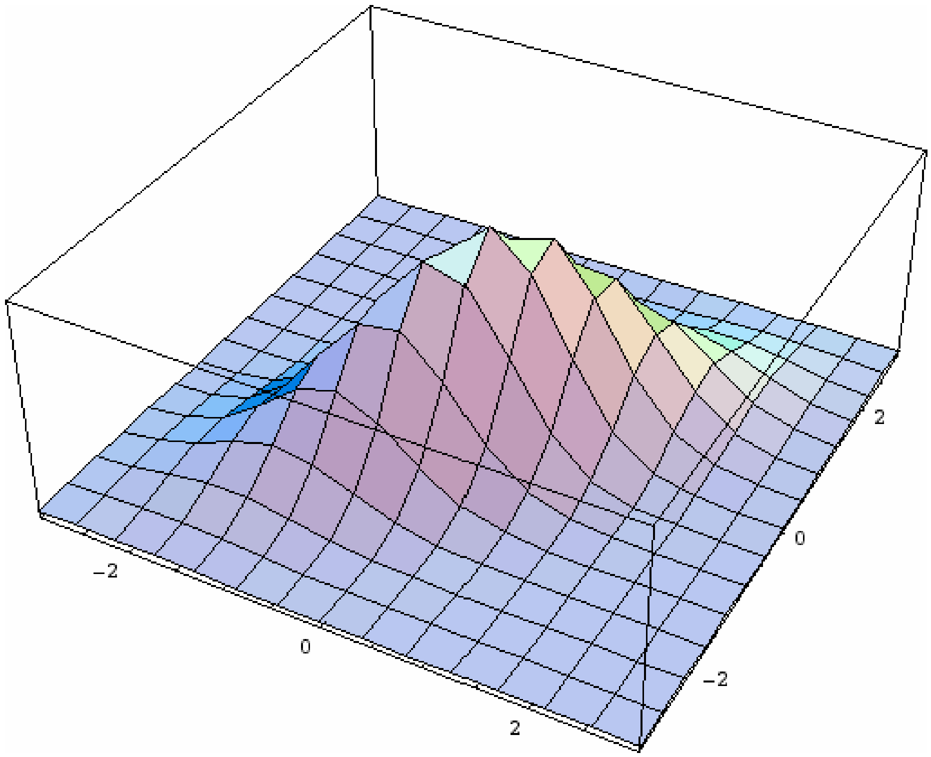

Figure 2. Example of an Elliptical Bivariate Normal Distribution as a Model for a Target Area

The process used to compute P_{TD} (Equation 1) depends on whether the density is specified to be a Uniform or Bivariate Normal distribution. These computations, which are discussed in detail in Gilbert et al. (2003a, pages 8 and 9), are conducted in the VSP software discussed below. First, VSP places the TA at a random location within the study site. Then, for the Uniform distribution, the survey transect design (pattern and spacing of transects) determined by VSP is laid over the site at a random starting position. If any portion of the TA is traversed by one or more transects, then the Binomial distribution is used to compute the probability that the geophysical sensor detects one or more of the anomalies that are expected to occur in the traversed portion of the TA based on the specified trigger and critical anomaly densities. This probability is computed using Equations 11 and 12 in Gilbert et. al. (2003a).

Next, the same survey transect design is again laid over the site starting at a new random position and the Binomial distribution is again used to compute the probability that the geophysical sensor detects one or more of the anomalies that are expected to occur in the traversed portion of the TA. This random process is repeated 10,000 times to obtain 10,000 estimates of $P_{D|T}$. The 10,000 estimates of $P_{D|T}$ differ from one another because the fraction of the randomly placed TA of specified size and shape that is traversed differs for each of the 10,000 iterations. The arithmetic average of these 10,000 estimates, denoted by $\overline{P}_{D|T}$, is then used in Equation (1) above to obtain the value of P_{TD} that is reported by VSP.

It should be noted that Gilbert et al. (2003a, page 8) indicate, as just explained, that the number of anomalies in the portion of a TA that is traversed by one or more transects randomly differ for

each of the 10,000 iterations. However, in order to reduce computation time, that method was changed. Now the VSP software uses the mathematical *expected* number of anomalies rather than the random number of anomalies for each of the 10,000 iterations, where the expected number in each iteration varies from that computed using the trigger density to that using the critical density of anomalies.

If the VSP user specifies that the anomaly density is a Bivariate Normal distribution, VSP computes the volume, V, of that distribution that lies along the one or more transects that transect the TA. This volume is used to compute the expected number of anomalies that lie in the transects that cross the TA when the anomaly density is the trigger anomaly density and then the critical anomaly density. Then Equations 11 and 12 in Gilbert et al. (2003a) are used to compute $P_{D|T}$ This process is repeated 10,000 times to obtain 10,000 values of $P_{D|T}$, which are averaged to obtain $\overline{P}_{D|T}$. This value of $\overline{P}_{D|T}$ is then used in Equation (1) above to compute P_{TD}.

PNNL incorporated its statistical design and data analysis methods for TA detection into the VSP software, which is being developed for environmental contamination characterization of soils, sediments, and surfaces in buildings (Hassig et al., 2002). [The report by Hassig et al, 2002, is currently being updated to apply to Version 3.0 of VSP. Version 3.0 and the new report should be available for download from http://dqo.pnl.gov/vsp within a few months.] The development of VSP is being co-sponsored by DOE, EPA, DoD, and DHS. VSP allows the user to import or draw a map of the site to be surveyed. Figure 3 shows a screen shot from the VSP UXO TA detection sampling module depicting a hypothetical site with the area where transect sampling is planned. Figures 4a and 4b illustrate the VSP dialog box that allows the user to input the DQOs listed above and other relevant design parameters. VSP calculates the required transect spacing to achieve the specified probability of traversing the critical TA of the specified size and shape (Figure 5), calculates the probability of both traversing and detecting the critical TA of the specified anomaly density if it exists, lays down the transects on the site map and outputs the x, y coordinates of the transects. Also, VSP automatically produces a report that summarizes the design and the outputs of VSP, which can be copied and pasted into project documents, technical papers, and publications.

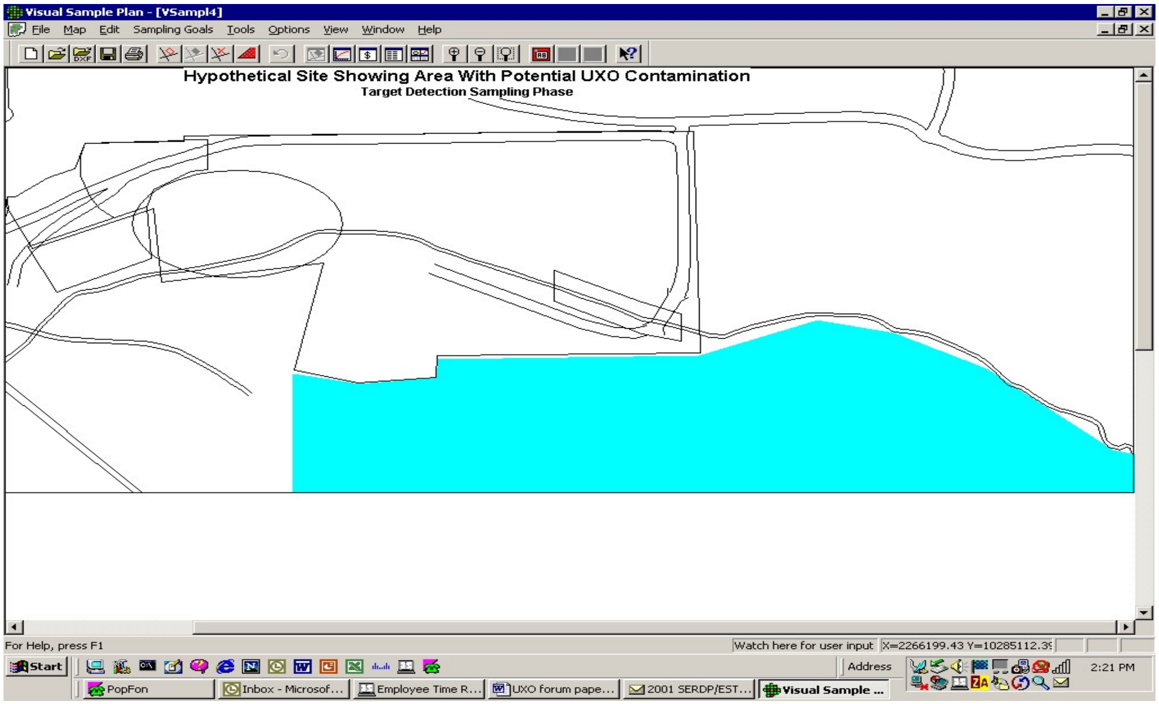

Figure 3. VSP Screen Shot Showing Site Map and Area to be Surveyed

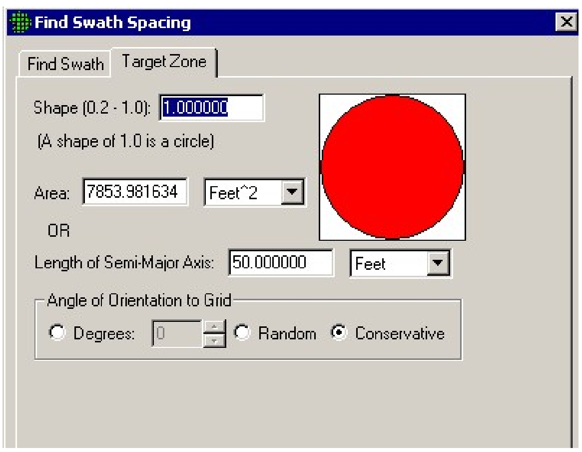

Figure 4a. VSP Dialog Box for Specifying the Target Area of Concern

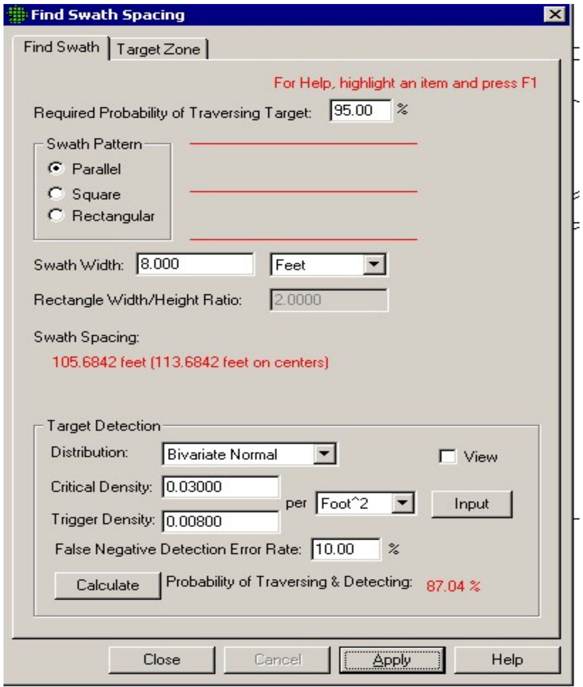

Figure 4b. VSP Dialog Box Showing DQO Inputs with Probability and Spacing Calculations

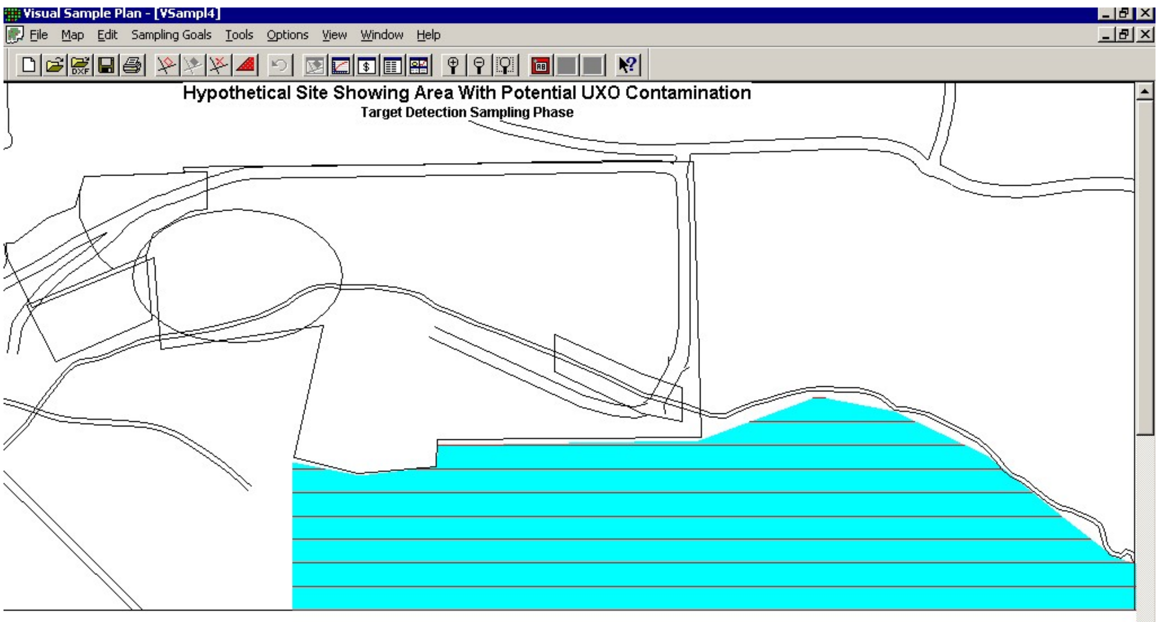

Figure 5. Site Map with Recommended Transect Locations

The VSP report includes a sensitivity analysis to show the effects of changing the VSP input parameters (DQOs) on the transect spacing required, percent coverage of the area, and cost.

These methods and tools allow for balancing DQO requirements against costs and other site constraints. If no TAs are identified, then plans for no further action might be considered. If potential TAs are identified by the geophysical survey, then the next phase of sampling may focus on defining the boundary of the TAs, developing a sufficiently accurate map of the anomaly/UXO density, and preparing a dig list.

4.1.2 Meandering Transects

The transect sampling plans in the previous section assume that perfect straight-line transects are spaced at regular intervals. However, when field crews implement a geophysical sensor survey, they may find that obstacles, dense vegetation, geographic features, and other factors make it undesirable, difficult, or impossible to follow straight lines with the geophysical sensors. Section 5.0 in Gilbert et al. (2003a) documents the methods developed by PNNL and used in VSP to approximate the probability that a TA of specified size and shape would be found using the straight-line or meandering transects that were actually used in the geophysical survey. Section 5.1 documents the methods developed to compute the probability that meandering transects traverse a TA. Sections 5.2 and 5.3 document the methods PNNL developed to compute the probability of both traversing and detecting the TA with meandering transects when the density of anomalies in the TA has a Uniform or Bivariate Normal distribution, respectively.

PNNL implemented in VSP the methods documented in Gilbert et al. (2003a) for computing the probability of traversing or of traversing and detecting an unknown TA of concern given a meandering or non-parallel characterization pattern. Figure 6 shows a combined meandering and somewhat parallel swathing pattern over the same site as previously illustrated in Figure 5. Simulations in VSP were used to determine that the probability is approximately 0.96 that the meandering pathway shown in Figure 6 would have traversed an elliptical target area that has a 100-foot major axis radius. The clusters of “Xs” shown on the map in Figure 6 are locations where a 100-foot elliptical TA could be present without being traversed (and hence not detected) by the meandering pathway.

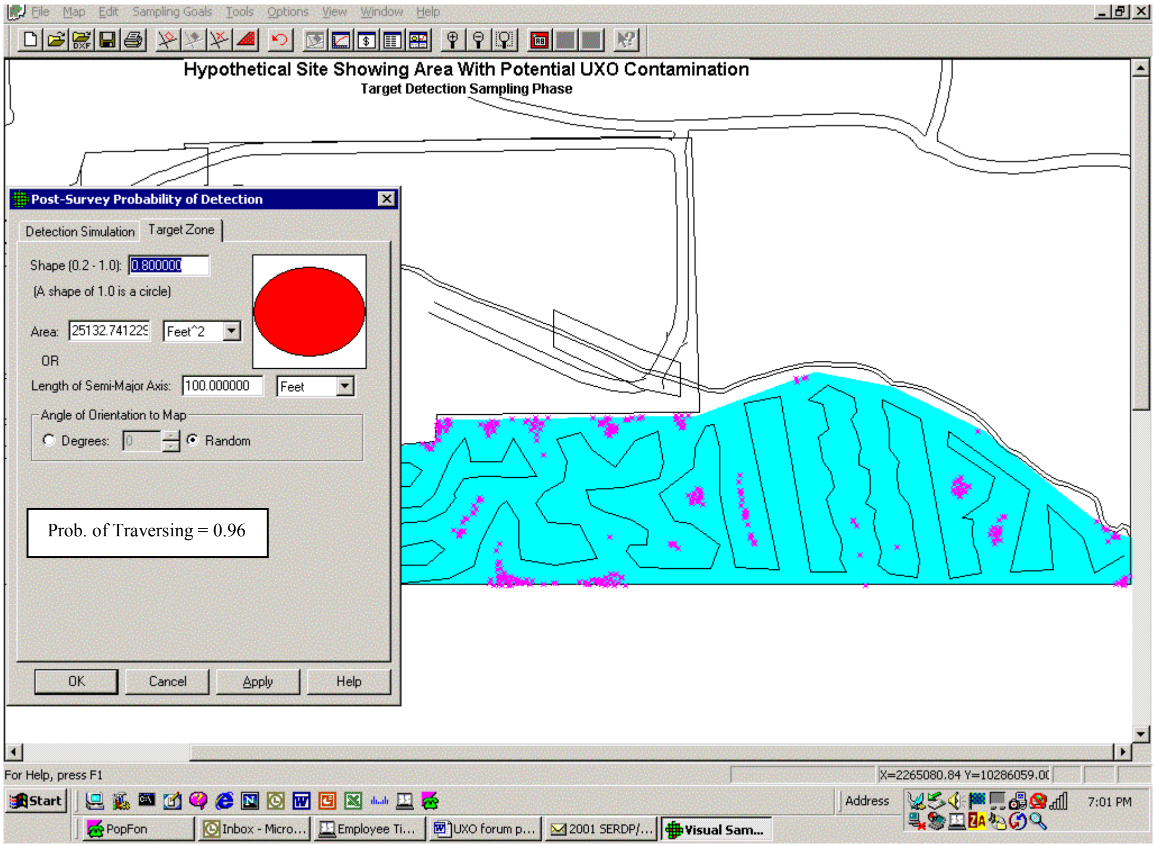

Figure 6. Meandering Pathway Performance Evaluation in VSP

4.2 Objective 2: Compliance Assessment

The compliance assessment objective was to develop a sampling strategy to achieve high confidence that no or very few UXO are present at a very large site or high confidence that no or few UXO remain at a site after UXO cleanup operations. After a review of the statistics literature PNNL selected two statistical approaches for incorporation into the VSP software tool: Schilling's method (Schilling 1982) and the Wright/Grieve approach (Wright 1992, Grieve 1994). These methods are described below.

4.2.1 Schilling's Method

Suppose there are N possible survey units (e.g., equal-length parallel transects or 100m by 100m plots) at a military base, but N is so large that only a fraction of the N units can be surveyed, given practical constraints. How should the number of survey units to survey, n, be determined? Schilling (1982) developed an acceptance sampling plan called "compliance sampling" for determining n that is widely used for industrial applications, e.g., by DoD (1999) to assess the acceptance of products in support of military standard MIL-STD-1916.

Schilling's method is implemented in the VSP software tool for when the survey unit is a transect (swath) and all transects have the same, or approximately the same, length. The procedure is as follows:

- The VSP user specifies the width of the transects (all transects have the same width).
- VSP uses the transect width and the map of the site to compute the total number, N, of potential parallel transects that could be surveyed in the study area.
- The VSP user specifies an upper limit, P_L, on the true fraction of the N transects that are "defective" that can be allowed. That is, if no more than 100 P_L percent of the N transects are "defective," then the site will be considered acceptable, although an "acceptable" site may still require certain actions to reduce risk from exposure to UXO.
- The VSP user specifies the 100(1-ε) percent confidence required that less than 100 P_L percent of the N transects are defective, where $0.01 < \varepsilon \leq 0.5$.
- VSP uses the method provided in Schilling (1982, pages 474 - 482) to compute the number, n, of the N transects that must be selected and geophysically surveyed, assuming zero false negative and false positive geophysical sensor error rates.
- If *none* of the n units that are surveyed are found to be defective, then it can be stated that the probability is 0.90 (90% confidence) that less than 100 P_L percent of the N transects are defective.

As explained in Schilling (1982, pages 474 - 482), n is determined by first computing D = N x P_L, where D is taken from a special table and used to find the fraction, f, of the N transects that must be surveyed. Then the number of transects is computed as n = f x N. The fractions, f, in Schilling's table are only appropriate when the required confidence is 90% (ε = 0.10). However, Schilling (1982, pages 478 - 479) shows how to determine the fraction, f, for any other confidence level desired. Verification that VSP is correctly computing n using Schilling's method is provided in Section 6.2 of Gilbert et al. (2003a).

Figure 7 illustrates Schilling's method for a simple, small rectangular study area that contains only N = 15 transects of equal width and length. In the VSP dialogue box the VSP user has specified that each transect is 5 feet wide and that 90% confidence is required that no more than 10 percent of the N = 15 transects are defective (that is, $P_L = 0.10$). VSP computes that n = 12 of the N = 15 transects must be surveyed and found to not be defective in order to make this confidence statement.

As indicated above, an assumption of Schilling's method is that the geophysical detector has a negligibly small false negative error rate (probability). Of course, this assumption is never fully applicable in practice. One possible approach to take into account the false negative error rate is to survey a larger number of the N transects than would be indicated under the assumption that false negative errors never occur. This approach can be implemented as follows:

Specify the best estimate of the probability, P_{fn}, that the geophysical sensor being used at the site will make a false negative error as it passes over a metal object of concern.

Compute P_{EL}, which is the "effective tolerable upper limit on the true fraction of the transects that are defective that can be allowed" (from Schilling 1982, page 85):

$$P_{EL} = P_L(1 - P_{fn})$$

Use P_{EL} rather than P_L in the dialog box for Schilling's method in the VSP software.

P_{EL} is always smaller than P_L because P_{fn} is always positive. This means that the number of transects that need to be surveyed and found not to be defective must be larger than when P_{fn} is assumed to be zero. That is, we compensate for the false negative error rate by surveying more transects in order to achieve the required confidence desired.

To illustrate, using the example above, suppose that the false negative error rate of the sensor is P_{fn} = 0.05 (5%), that no more than 10 percent of the N transects are "defective," i.e., P_L = 0.10. and we desire 90% confidence. First, compute

$$P_{EL} = P_L(1 - P_{fn}) = 0.10\ (1 - 0.05) = 0.95$$

If we enter P_{EL} = 9.5 % rather than P_L = 10 % into the VSP dialogue box for Schilling's method (Figure 7), VSP computes that n = 13 rather than 12 transects should be surveyed. If all 13 transects are not defective, then we have 90% confidence that the percentage of defective transects does not exceed 100 P_L = 10%.

Schilling's method also assumes that there is a negligibly small false positive error rate. It is possible to compute P_{EL} when both false positive and false negative error rates are greater than zero (see Schilling 1982, page 85). However, a non-zero false positive rate has the effect of reducing the number of transects to be surveyed in order to compensate for false positive detections. This non-conservative approach does not seem wise. False positive errors increase cost by increasing the number of locations that need to be dug or investigated, but at least false positive errors can, in that way, be identified. Of course, improvements in the discrimination ability of geophysical sensors would reduce the number of false positive errors.

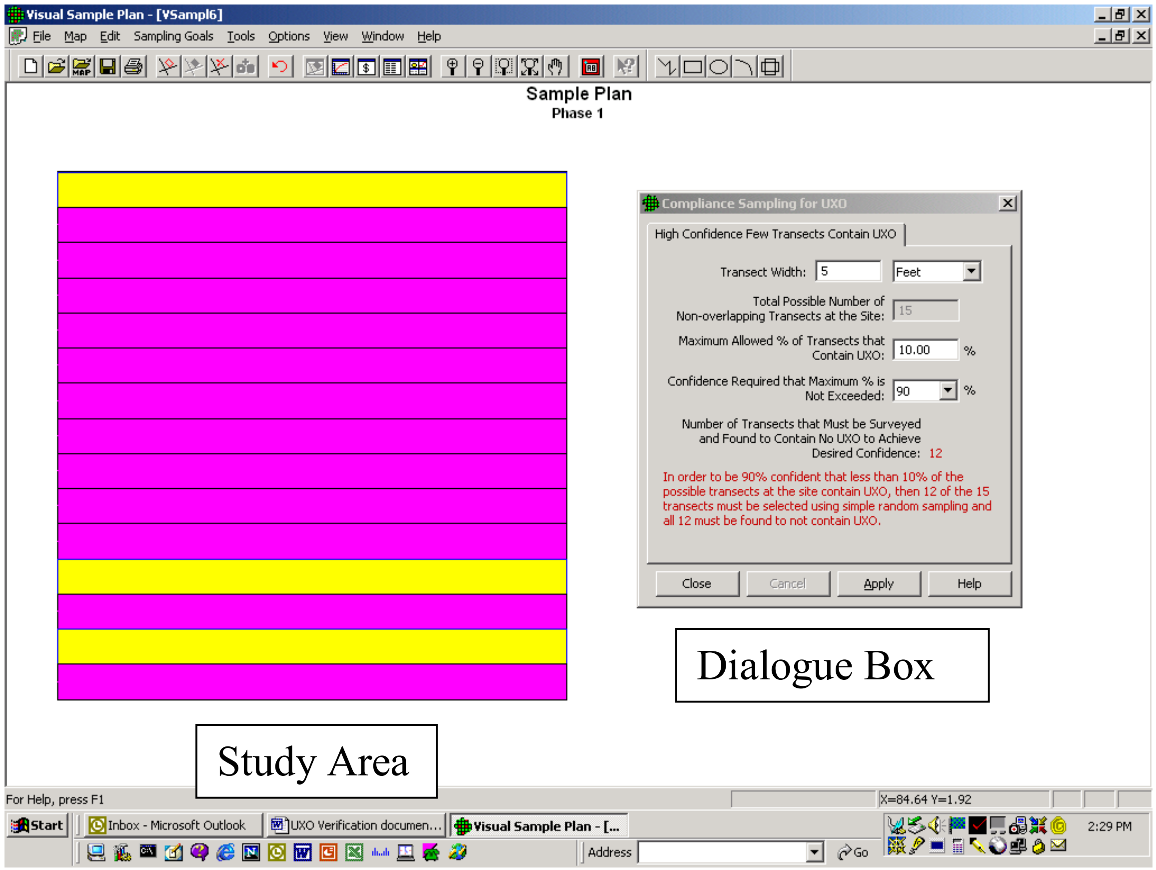

Figure 7. Illustration of Schilling's Method (from VSP Software)

4.2.2 Wright/Grieve Method

Grieve (1994) expanded on the method in Wright (1992) to develop an equation (his Equation 2.5) that can be used to compute the number of survey units (e.g., transects), n, that should be selected from the total set of N transects and found to be non-defective in order to be 100(1-ε) percent confident that all transects not surveyed are also not defective. Grieve's equation is used in the VSP software.

Wright and Grieve's method is "Bayesian" because it requires that the stakeholders provide a quantitative measure of their *belief* that the study area contains no defective survey units. This belief should be based on all information and data collected about the study area and the conceptual site model (CSM) developed for the area. The stakeholders quantify their "belief" by choosing a specific Beta probability distribution for the fraction, f, of the N units that are defective. [The Beta distribution is described in, e.g., Rothschild and Logothetis (1986, pages 50-51), Patel et al (1976) and Johnson and Kotz (1970).] In other words, the stakeholders are uncertain about the fraction, f, of transects that are defective, but they can agree that the probability that this fraction takes on various values can be modeled by a specific Beta distribution, the shape of which is determined by the value of the two parameters of the distribution: a and b. The expected (true average) value, δ, of f for a Beta distribution with parameter values a and b is $\delta = a/(a+b)$.

The VSP software allows the VSP user to choose among 7 possible Beta distributions. These distributions are listed in Table 1.0 and are illustrated in Figures 8-14. The shape of each distribution and the expected value δ for each distribution is determined by the values of the two parameters, a and b.

Table 1. The Seven Beta Distributions Available for Selection in VSP Software to Model the Uncertainty in the Fraction of Transects that are Defective

Parameter Values of the Seven Beta Distributions in VSP	Expected Value, δ, of the Fraction, f, of Transects that are Defective*	English Characterization of the Beta Distribution used in the VSP Software
1. a = 1, b = 999	0.001	Extremely low fraction
2. a = 1, b = 99	0.01	Very low fraction
3. a = 1, b = 9	0.1	Low fraction
4. a = 1, b = 1	0.5	All fractions equally likely
5. a = 9, b = 1	0.9	High fraction
6. a = 99, b = 1	0.99	Very high fraction
7. a= 999, b = 1	0.999	Extremely high fraction

* δ = a/(a+b)

The VSP user selects one of the seven distributions and VSP computes n using the following equation (derived from Equation 2.5 in Grieve 1994):

$$n \geq N - (N+b)\left\{1-(1-\varepsilon)^{(1-\delta)/(b\delta)}\right\} \qquad (2)$$

where N, a, b, δ, and 1-ε have been defined above. If the geophysical sensor survey does not find any defective survey units among the n randomly selected transects, then one can state with 100(1-ε) percent confidence that there are also no defective units in any of the N-n transects that were not surveyed. As is the case for Schilling's method, it is assumed that all survey units (e.g., transects) are equal in size.

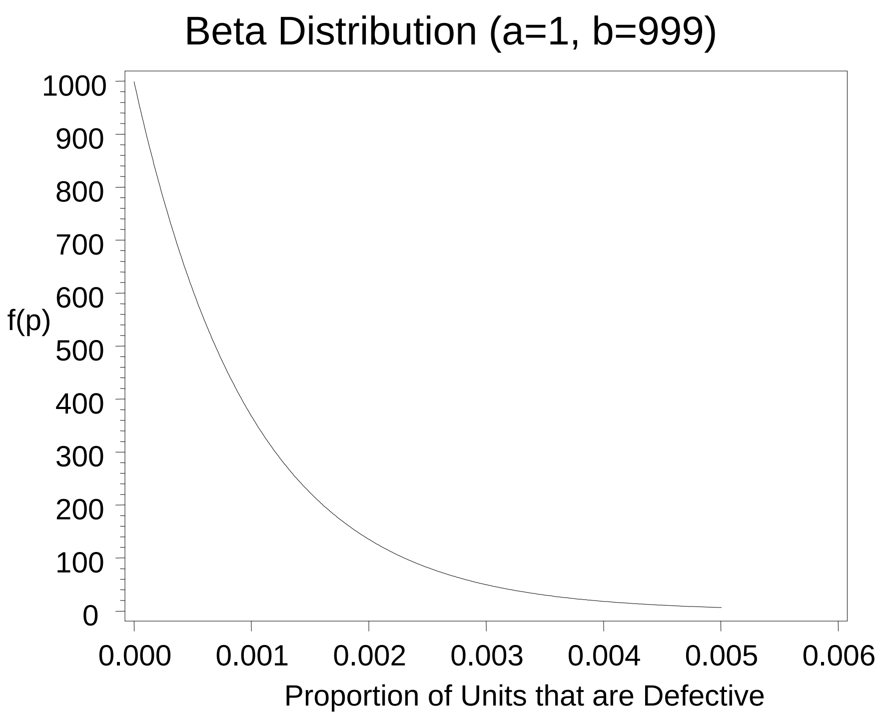

Figure 8. Beta Distribution with Parameters a = 1, b = 999 and Expected δ = 0.001

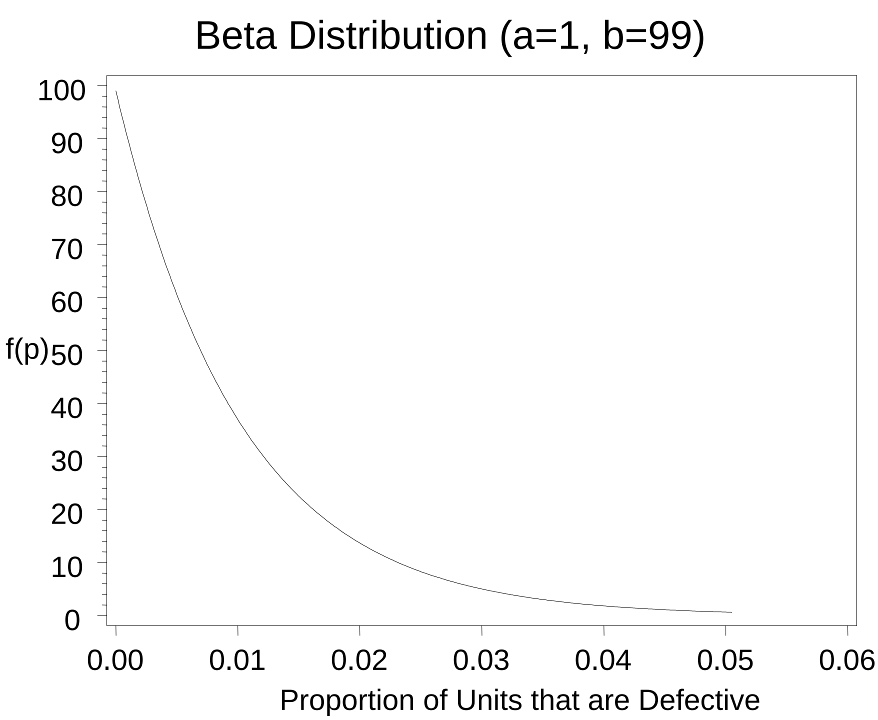

Figure 9. Beta Distribution with Parameters a = 1, b = 99 and Expected Value δ = 0.01

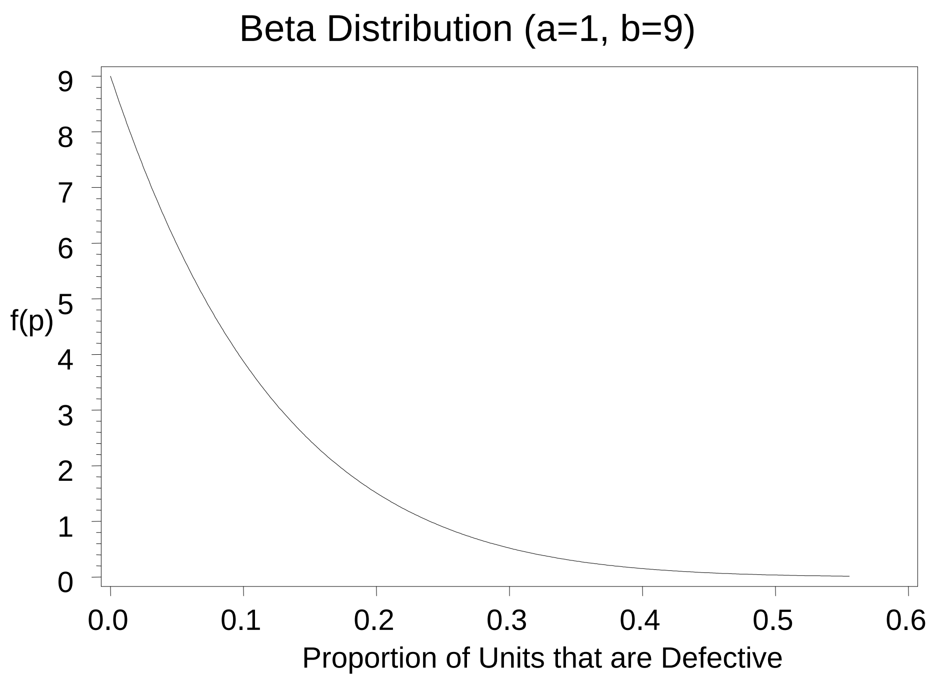

Figure10. Beta Distribution with Parameters a = 1, b = 9 and Expected Value δ = 0.1

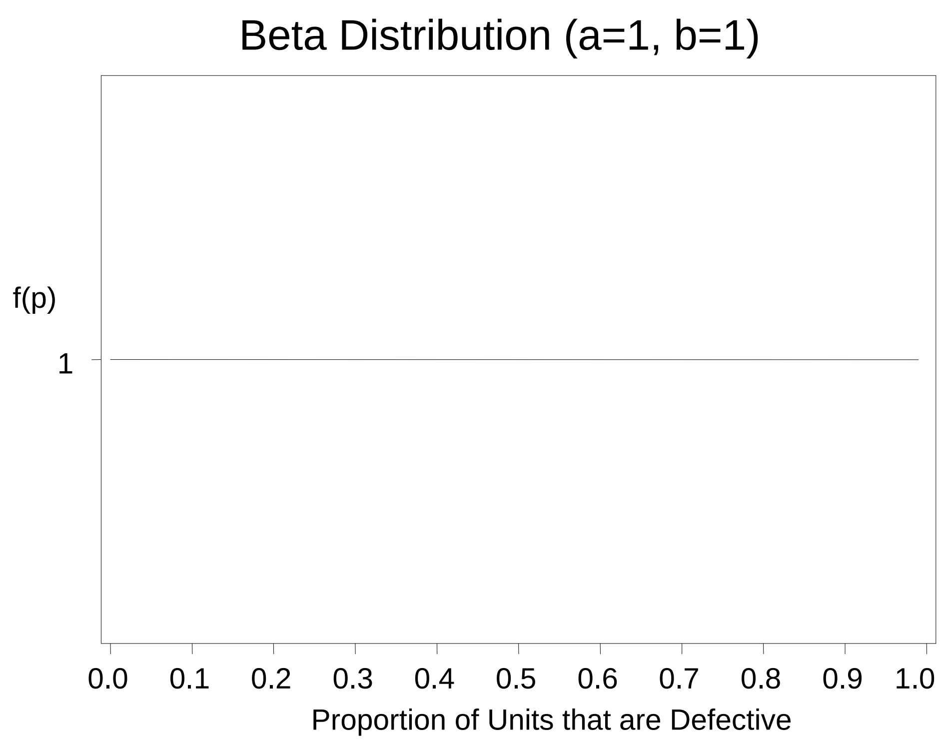

Figure 11. Beta Distribution with Parameters a = 1, b = 1 and Expected Value δ = 0.5

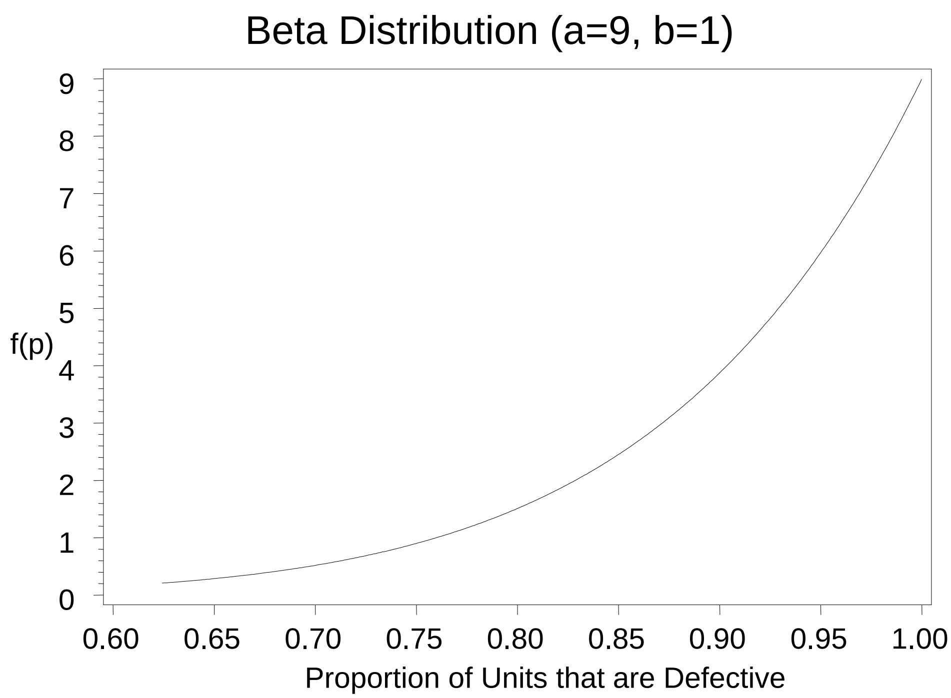

Figure 12. Beta Distribution with Parameters a = 9, b = 1 and Expected Value δ = 0.9

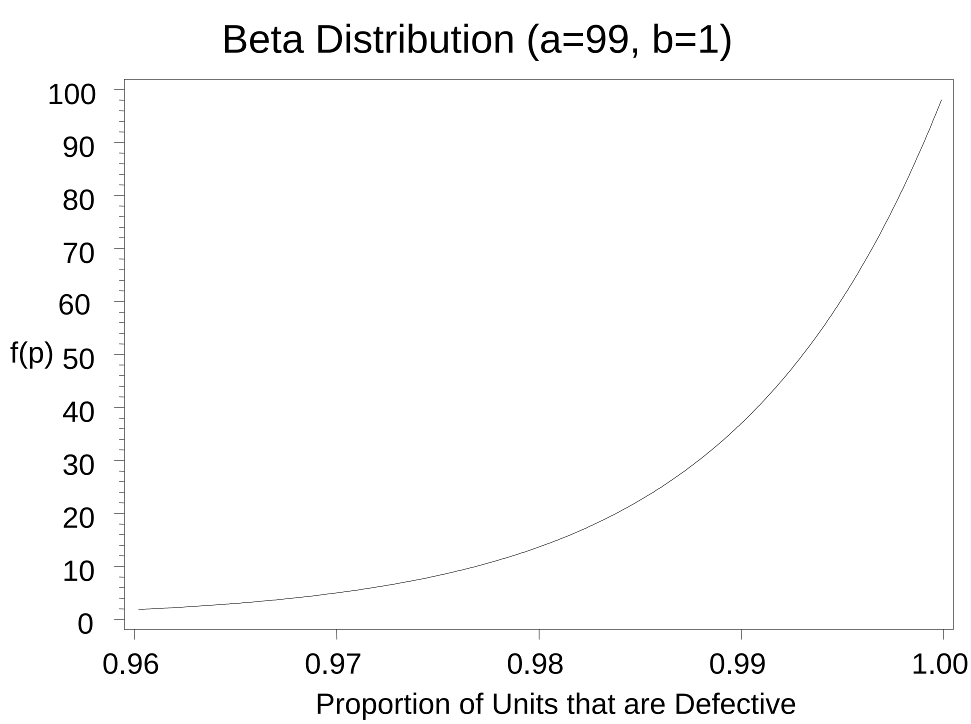

Figure 13. Beta Distribution with Parameters a = 99, b = 1 and Expected Value δ = 0.99

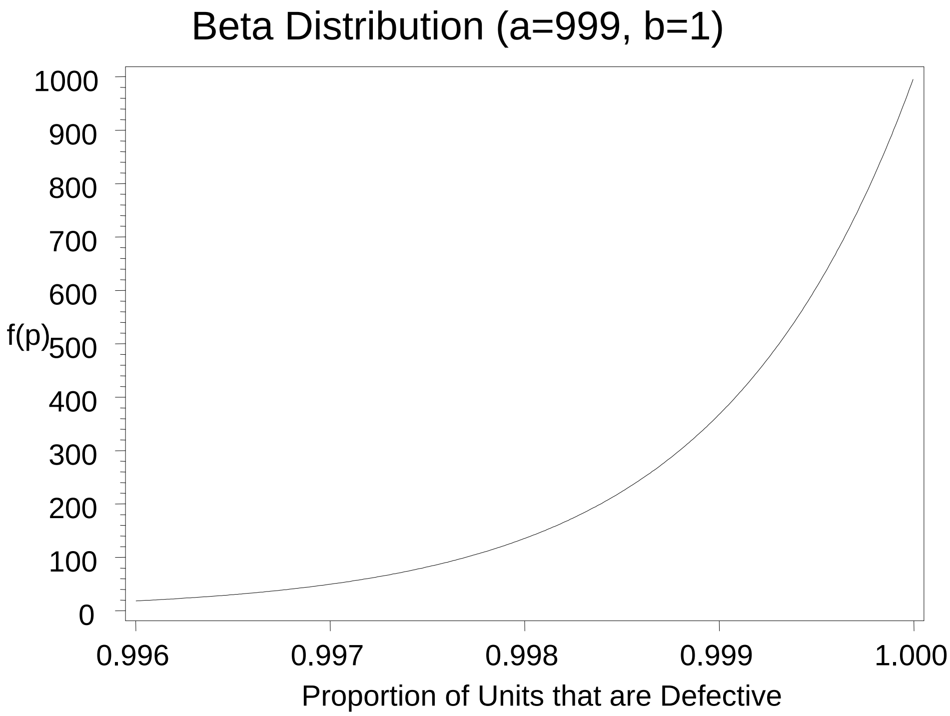

Figure 14. Beta Distribution with Parameters a = 999, b = 1 and Expected Value δ = 0.999

Figure 15 illustrates the Wright/Grieve method. A rectangular study area is shown; the same one as shown in Figure 7 for Schilling's method. In the dialogue box the VSP user has specified that each transect (survey unit) is 5 feet wide, that he/she believes that 10% of the transects in the study area may be defective, and that 95% confidence is required that there are no defective units in the study area if no defective units are found in the n transects surveyed. With these specifications, VSP computes that 14 of the 15 transects must be surveyed and found to not be defective in order to have 95% confidence that no transects in the study area are defective, i.e., that the single remaining un-surveyed transect is also not defective.

As was the case for Schilling's method, the Wright/Grieve method assumes that the geophysical detector used in the field has negligibly small false negative and false positive error rates. However, no adjustments to the number of survey units that should be surveyed to account for sensor errors have been developed by this PNNL project. Additional research is recommended.

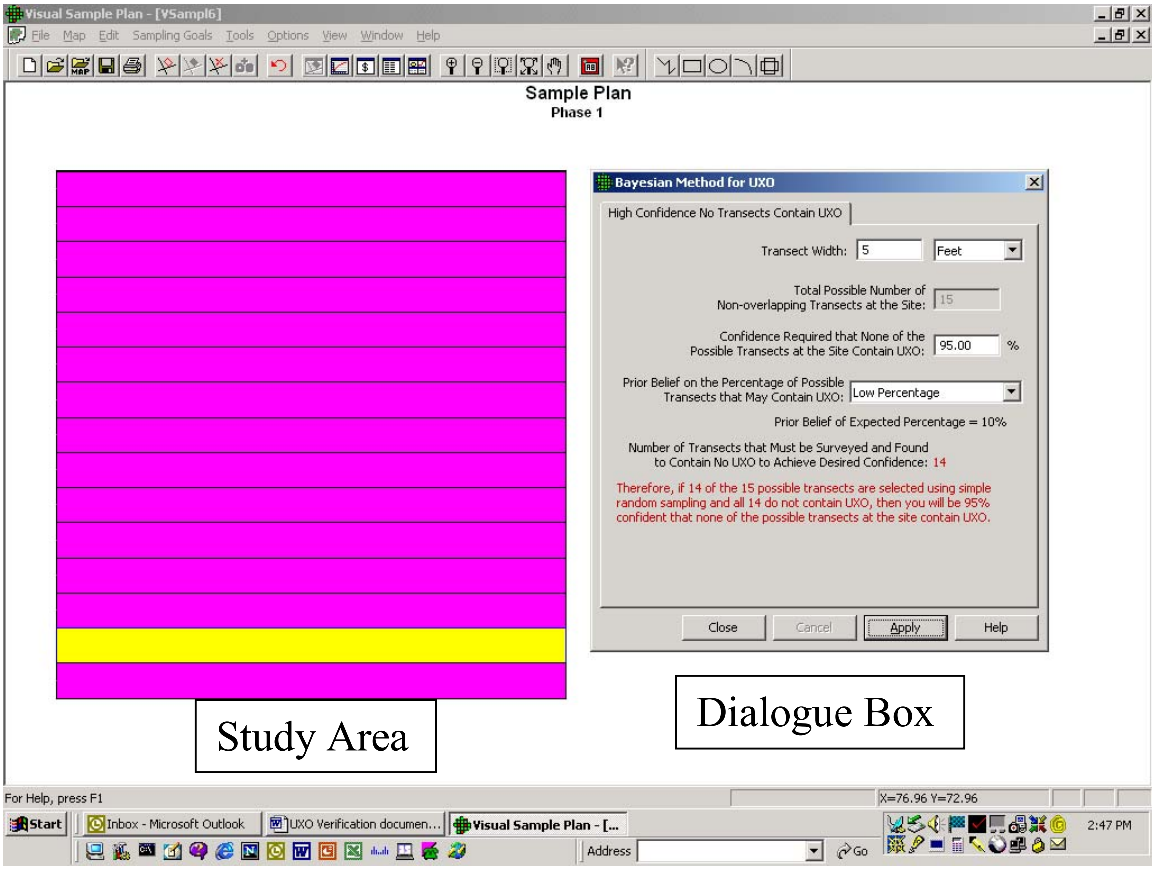

Figure 15. Illustration of the Wright-Grieve Method in VSP for Determining the Number of Transects (Survey Units) to Survey

4.3 Objective 3: Dig List Stopping Rules

The problem addressed by PNNL was to develop statistical methodologies to determine an early stopping rule such that, after n sample dig locations ($n < N$) have been observed, the upper bound on the probability that there is any UXO left in the remaining $N - n$ un-sampled locations is small. (Note that in this section of the report the notations N and n denote the number of specific field *locations*, whereas in Section 3.0 above they denote the number of potential and surveyed survey units (e.g., transects in the area.) PNNL developed a Bayesian statistical formulation to the problem and proposed a Bayesian estimator and bound which account for the site-dependent probability that a possible UXO location actually contains UXO.

4.3.1 Proposed Bayesian Methodologies

Wright (1992) gives a stopping rule when each location has the same probability of containing UXO and when no UXO has been found in an initial sample of n locations. However, this methodology does not take into account that the N locations have been characterized from "most likely" to "least likely" to be UXO based on a geophysical survey and expert judgment. The probability that a location contains UXO changes depending on which bin that location is assigned. Further, as shown below, the probability that a location contains UXO also depends on

the true proportion of UXO on the site. This site-dependent UXO proportion is a "prior probability" that can be estimated from the results of the initial samples.

A Bayesian formulation takes into account the binning scheme of the locations and the site-dependent prior probability in estimating the probability that the remaining un-sampled locations contain any UXO. The Bayesian methodology described here puts no restrictions on the number of UXO found in the initial sample. Therefore, the method is applicable if 0, 1, 2 or more UXO are found in the initial sample.

The Bayesian formulation assumes the following for the site of interest:

1. The performance of the binning scheme at the site is available, that is, the probabilities of UXO and non-UXO locations being assigned into each of the *k* bins are fixed (estimated from instrument performance studies or experience from a similar site),
2. Without knowing their bin assignments, the locations at the site of interest have the same prior probability of containing UXO,
3. Any spatial trends for UXO distribution across the site are incorporated into the binning scheme.

The performance probabilities of Assumption 1 are put into a performance matrix. The performance matrix can be estimated using real calibration data such as that in McDonald et al. (2001). It is reasonable to assume that the estimated performance matrix is appropriate for similar sites. If not, it would need to be adjusted for the particular site of interest.

In Section 4.3.2 below a Bayesian methodology is proposed that takes into account the geophysical measurement and expert judgment characterizations and the performance matrix. In Section 4.3.3, Bayes Rule (Casella and Berger 1990) is used to calculate the probabilities that the unobserved locations contain UXO. These individual probabilities are then combined to give the probability that the site contains any UXO. Section 4.3.4 contains a simple two-bin illustration of the methodology and Section 4.3.5 contains a more realistic six-bin example. Section 4.3.6 provides a summary and conclusions.

4.3.2 Mathematical Formulation of the Problem

Assume that there are *N* suspected UXO locations on the initial dig list for the site of interest. The true state of each location is given by X_i, which is a random variable defined as

$$X_i = \begin{cases} 1 & \text{if UXO is present} \\ 0 & \text{if UXO is not present} \end{cases} \quad \text{for } i = 1, 2, \ldots, N. \tag{3}$$

Before the geophysical measurements have been taken, each X_i can be considered to be a random sample from a Bernoulli distribution with probability determined by its expectation:

$$E(X_i) = Pr(X_i = 1) = p, \tag{4}$$

where p is the proportion of identified locations on the site that contain UXO. Of course, this p is unknown until all locations have been dug, but it is an important quantity we want to estimate.

Let Y_i represent the binning response from the geophysical measurement system at the i^{th} location. The response Y_i is a discrete random variable with k possible outcomes ordered from 1 to k, where 1 means "most likely" to contain UXO and k means "least likely" to contain UXO. The conditional probability distribution of the response Y_i given the true state X_i is determined by the geophysical measurement system performance matrix, which is denoted here as Matrix A. That is, there is a ($k \times 2$) matrix A with (j, l)-element A_{jl}, such that

$$A_{jl} = \Pr\left(Y_i = j \middle| X_i = l\right) \qquad \text{for } j = 1, 2, \ldots, k \text{ and } l = 0, 1 \tag{5}$$

This A matrix is shown in Table 2.

Table 2. Performance Matrix A for the Bayesian Dig List Method

Bin	Probability that a Location that is Truly ***Not*** UXO is Assigned to this Bin	Probability that a Location that is Truly UXO is Assigned to this Bin
1 - Highest Likelihood of UXO	A_{10}	A_{11}
2 - High Likelihood of UXO	A_{20}	A_{21}
$\vdots$	$\vdots$	$\vdots$
k - Lowest Likelihood of UXO	A_{k0}	A_{k1}

The problem is to use the geophysical measurement system responses to identify a set of n locations to be dug up and then, based on what is observed, to estimate the probability that there are no UXO in the N - n remaining locations. If this probability, P_0, is large enough, then digging can be stopped. The risk of any UXO remaining in the un-sampled dig list locations is 1 - P_0. Mathematically, P_0 is represented as follows.

Let $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_N)$ be the vector of binning responses that is based on the geophysical measurement system; N_j of which are equal to j, for $j = 1, 2, \ldots, k$. Let D be the set of n indices of the locations that are dug up and R be the set of indices of the remaining locations (so that $D \cup R = \{1, 2, \ldots, N\}$). The probability P_0 of no UXO in the remaining locations given the available information is the conditional probability

$$P_0 = \Pr\left(\sum_{j \in R} X_j = 0 \middle| \mathbf{Y} \text{ and } \{X_i ; i \in D\}\right). \tag{6}$$

The probability that any of the remaining locations contain UXO is 1 - P_0.

4.3.3 Estimating and Bounding P_0

A first step in estimating P_0 is to use Bayes Rule to determine the distribution of X_i given Y_i from matrix A and p (Equation 4). That is, explicitly,

$$
\begin{aligned}
\Pr\left(X_i = 0 \mid Y_i = j\right) &= \frac{\Pr\left(X_i = 0 \text{ and } Y_i = j\right)}{Pr\left(Y_i = j\right)} \\
&= \frac{\Pr\left(Y_i = j \text{ and } X_i = 0\right)}{\Pr\left(Y_i = j \text{ and } X_i = 0\right) + \Pr\left(Y_i = j \text{ and } X_i = 1\right)} \\
&= \frac{\Pr\left(Y_i = j \mid X_i = 0\right)\Pr\left(X_i = 0\right)}{\Pr\left(Y_i = j \mid X_i = 0\right)\Pr\left(X_i = 0\right) + \Pr\left(Y_i = j \mid X_i = 1\right)\Pr\left(X_i = 1\right)} \\
&= \frac{A_{j0}(1-p)}{A_{j0}(1-p) + A_{j1}p}
\end{aligned}
\tag{7}
$$

and similarly,

$$
\Pr\left(X_i = 1 \mid Y_i = j\right) = \frac{A_{j1}p}{A_{j0}(1-p) + A_{j1}p}. \tag{8}
$$

We estimate P_0 by first using the bin information $\mathbf{Y}$ and the dig information $\{X_j; j \in D\}$ to obtain an estimate, $\hat{p}$, of the site UXO proportion p, and then estimate

$$
\hat{P}_0 = \coprod_{i \in R} \hat{\Pr}\left(X_i = 0 \mid Y_i = j\right), \tag{9}
$$

Then, also use $\hat{p}$ in Equation (7) to obtain

$$
\hat{\Pr}\left(X_i = 0 \mid Y_i = j\right) = \frac{A_{j0}(1-\hat{p})}{A_{j0}(1-\hat{p}) + A_{j1}\hat{p}}. \tag{10}
$$

The estimation of p is Bayesian. In general, a Bayesian analysis provides a framework for assessing the information regarding the parameters that is contained in the available data. In this case, the parameter is p and the data are the binning information and the results from the locations that have been dug up. The information is expressed in terms of the posterior conditional density function of $p \in (0, 1)$ given $\mathbf{Y}$ and $\{X_i; i \in D\}$. Suppose that the set of indices D corresponds to the N_1 bin 1 locations, those that the geophysical measurement system gives the highest likelihood of containing UXO. Then the posterior density function is

$$
f\left(p \middle| \mathbf{Y} \textit{ and } \{X_i; i \in D\}\right) \propto \left(\frac{A_{10}(1-p)}{A_{11}p + A_{10}(1-p)}\right)^{n_{10}} \left(\frac{A_{11}p}{A_{11}p + A_{10}(1-p)}\right)^{n_{11}} g\left(p \middle| \mathbf{Y}\right), \tag{11}
$$

where

$$
g\left(p \middle| Y\right) = B(p; \alpha, \beta) \prod_{j=1}^{k} \left(A_{j1}p + A_{j0}(1-p)\right)^{N_j}, \tag{12}
$$

n_{10} and n_{11} are the number of $Y_i = 1$ locations which have $X_i = 0$ and $X_i = 1$, respectively, $n_{10} + n_{11} = N_1$, and $B(p;\ \alpha, \beta)$ is the prior Beta distribution density function with parameters α and β,

$$B(p;\alpha,\beta)=\begin{cases}\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1} & \text{for } 0<p<1\\ 0 & \text{otherwise}\end{cases}. \tag{13}$$

The constant of proportionality in Equation (11) is determined by numerical integration so that the integral of $f(\cdot)$ over (0, 1) equals one. The function $g(\cdot)$, properly normalized, gives the posterior density for *p* given **Y** before digging up any locations.

The family of Beta distributions is extremely flexible. For instance, a uniform distribution, typically chosen to reflect a non-informative prior, is obtained by setting $\alpha = \beta = 1$. Analysis using the uniform distribution produces the Maximum Likelihood Estimation (MLE) of *p*. Choice of α and β for real prior information is made easy using the expressions for the mean and variance of the Beta distribution:

$$\mu=\frac{\alpha}{\alpha+\beta} \text{ and } \sigma^2=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \tag{14}$$

So if prior information on *p* is represented as a mean μ and a variance σ^2, then Equation (14) provides the values of α and β for the Beta prior distribution in Equations (12) and (13).

An estimate of *p* is the value $\hat{p} \in$ (0, 1) that maximizes the posterior density function f (p $\mid$ **Y** and { X_j; *j D*). If $\alpha = \beta = 1$ then $\hat{p}$ is the Maximum Likelihood Estimate (MLE). A 95% confidence upper bound, $\hat{p}_{95}$, can be found such that

$$\int_0^{\hat{p}_{95}} f\left(p \mid \mathbf{Y} \text{ and } \{X_j; j \in D\}\right)dp = 0.95. \tag{15}$$

This upper bound $\hat{p}_{95}$ is used to determine a 95% confidence lower bound for P_0 through Equations (9) and (10). A lower bound for P_0 is obtained because the probabilities in Equation (7) are monotone decreasing functions of *p*.

Information from other locations is accounted for by multiplying the right-hand side of Equation (11) with terms like

$$\left(\frac{A_{j0}(1-p)}{A_{j1}p+A_{j0}(1-p)}\right)^{n_{j0}}\left(\frac{A_{j1}p}{A_{j1}p+A_{j0}(1-p)}\right)^{n_{j1}} \tag{16}$$

where n_{j0} and n_{j1} are the number of bin *j* locations dug up and observed to have $X_i = 0$ (not UXO) and $X_i = 1$ (UXO), respectively.

4.3.4 Two-Bin Examples

This section presents example numerical results of the methodology in Sections 4.3.1, 4.3.2, and 4.3.3. These examples are not realistic, but are presented to illustrate in the simplest terms the mathematical approach. We consider a geophysical measurement system with responses in just *k*

= 2 bins. For a first example, consider the hypothetical results of a calibration study given in Table 3. Such a study would consist of seeding a site with UXO and non-UXO and running the geophysical measurement system over it.

Table 3. Two Bin Example 1: Calibration Data Matrix

Bin	# Not UXO	# UXO
1 - High Likelihood of UXO	10	76
2 - Low Likelihood of UXO	90	4
Totals	100	80

From Table 3.0 we can calculate two sets of conditioned probabilities. These two sets are:

Set 1:

Probability of being placed in high likelihood UXO bin | It is UXO = 0.95
Probability of being placed in low likelihood UXO bin | It is not UXO = 0.90
Probability of being placed in high likelihood UXO bin | It is not UXO = 0.10 (false positive)
Probability of being placed in low likelihood UXO bin | It is UXO = 0.05 (false negative)

Set 2:

Probability of being UXO | It was placed in high likelihood UXO bin = 0.88
Probability of being non UXO | It was placed in low likelihood UXO bin = 0.96
Probability of Not being UXO | It was placed in high likelihood UXO bin = 0.12
Probability of Not being UXO | It was placed in low likelihood UXO bin = 0.04

The probabilities in Set 2 are completely dependent on how many UXO and non-UXO were seeded in the calibration experiment and do not reflect the performance of the binning procedure. Moreover they are variant depending on the calibration site setup.

However, the probabilities in Set 1 are invariant and would be expected to remain fairly consistent across several calibration sites. The probabilities in Set 1 measure the performance of the binning procedure and provide the Prob($Y_i = j \mid X_i = l$) = A_{jl} needed for the Bayesian formulation shown in Equation (8). The resulting estimated A matrix is given in Table 4.

Table 4: Two Bin Example 1: Matrix A

Bin	Probability that a Location that is Truly ***Not*** UXO is Assigned to this Bin	Probability that a Location that is Truly UXO is Assigned to this Bin
1 - High Likelihood of UXO	0.10	0.95
2 - Low Likelihood of UXO	0.90	0.05

Assume that 200 locations were identified at a new site and that p = 10% of them contain UXO. Data were simulated using the model of Section 4.3.2 (following the random distributions in Equations 3-5. The binning yielded 33 bin 1 locations and 167 bin 2 locations (see Table 5). Three stages of analysis were considered to illustrate the methodology in Sections 4.3.1, 4.3.2 and 4.3.3. The first stage was to dig up the bin 1 locations. The resulting UXO/Not UXO counts for

each stage are given in Table 6. The second stage was to sample 67 of the bin 2 locations to bring the total to 100 locations dug up. The third stage was to sample 50 more of the bin 2 locations to bring the total to 150 locations dug up. A uniform (noninformative) prior on p (that is, a Beta (α = 1, β = 1) distribution) was used to in the analysis to estimate $\hat{P}_0$ and the 95% lower bound (using the 95% upper bound for p) after each stage of sampling; see the results in Table 6. After the third stage, the probability of no UXO in the 50 remaining locations was estimated to be > 64.3% (with 95% confidence). Digging up more locations would be required to reduce the risk of leaving UXO at the site.

Table 5. Two Bin Example 1: Simulated Data

Bin	Counts N_j	First Stage Not UXO	First Stage UXO	Second Stage Not UXO	Second Stage UXO	Third Stage Not UXO	Third Stage UXO
1 - High Likelihood	33	14	19	14	19	14	19
2 - Low Likelihood	167			67	0	117	0

Table 6. Two Bin Example 1: Analysis Results

Stage	Number Dug Up	Total Dug Up	$\hat{P}_0$	Probability of No UXO Remaining 95% Confidence Lower Bound
First	33	33	0.357	0.214
Second	67	100	0.548	0.407
Third	50	150	0.744	0.643

The second two-bin example used the "better" A matrix given in Table 7. A lower false negative rate (the probability of assigning a UXO location to the "Low Likelihood" bin) of 1% was considered, compared to 5% in the first example. Using the same p = 10% and noninformative prior, the data in Table 8 were simulated and the results in Table 9 were calculated. After the third stage, the probability of no UXO in the 50 remaining locations was estimated to be > 91.6% (with 95% confidence). This is much higher than the probability from the first example.

The limitations of these examples are evident and can be attributed to the oversimplification of binning into only two bins. The next section illustrates the practicality and potential for justifying no further digs when the binning procedure involves a more realistic 6 bins.

Table 7. Two Bin Example 2: Matrix A

Bin	Probability that a Location that is Truly ***Not*** UXO is Assigned to this Bin	Probability that a Location that is Truly UXO is Assigned to this Bin
1 - High Likelihood of UXO	0.10	0.99
2 - Low Likelihood of UXO	0.90	0.01

Table 8. Two Bin Example 2: Simulated Data

Bin	Counts N_j	First Stage		Second Stage		Third Stage	
		Not UXO	UXO	Not UXO	UXO	Not UXO	UXO
1 - High Likelihood	31	13	18	13	18	13	18
2 - Low Likelihood	169			68	1	118	1

Table 9. Two Bin Example 2: Analysis Results

Stage	# Dug Up	Total Dug Up	$\hat{P}_0$	95% Confidence Lower Bound for P_0
First	31	31	0.829	0.753
Second	69	100	0.889	0.839
Third	50	150	0.943	0.916

4.3.5 Six-Bin Methodology Using Badlands Performance Matrix

This section presents example numerical results of the methodology in Sections 4.3.1, 4.3.2, and 4.3.3 applied to a six-bin scheme. The performance matrix used for the calculations in this section is based on a study performed on the Badlands Bombing Range (BBR) discussed in McDonald et al. (2001). Data from this study were used to create a performance matrix, *A*, so that this example used here had some basis in reality. Some modifications were made to the data set such as rounding the count of dig locations and interpretation of the study results to form the performance matrix A. This was done because the objective of the McDonald et al. study was to compare the results of two technologies used to detect UXO, not to determine the performance matrix A.

For this example, a 6 category binning strategy is illustrated, i.e., anomaly locations are placed into 6 categories. Data from McDonald et al. (2001), which describes a study of a Multi-Sensor Towed Array at the BBR, was used to construct the performance matrix A for this binning strategy. This performance matrix was formed using the BBR data, with some modifications, and will be referred to herein as the “Test” A matrix. This matrix was formed solely to perform the calculations described in this paper to demonstrate the stopping rule methodology, so that the results had some basis in reality. This “Test” A matrix is in no manner an analysis or statement concerning the discrimination capability used in the BBR study.

The “Test A” matrix was formed from the BBR data from observations recovered from a 100% analysis of anomalies in the “Seed Target Area” that was seeded with inert UXO (page 47 of McDonald et al. 2001). The count of UXO and non-UXO was first put into a matrix shown in Table 10.

Table 10. Badlands Bombing Range Test Results for the Bayesian Dig List Method

Category	Count of Locations Truly Not UXO	Count of Locations Truly UXO
1 –High	6	18
2	10	5
3	36	0
4	3	0
5	36	1
6 – Low	54	1
Total	145	25

The conditional probabilities were calculated as shown in Table 11.

Table 11. Conditional Probabilities for Badlands Bombing Range Test Results

Category	Probability Assigned to Category i Given that the Location is Truly Not UXO	Probability Assigned to Category i Given that the Location is Truly UXO
1 -High	6/145 = 0.04	18/25 = 0.720
2	10/145 = 0.07	5/25 = 0.200
3	36/145 = 0.25	0/25 = 0.000
4	3/145 = 0.02	0/25 = 0.000
5	36/145 = 0.25	1/25 = 0.040
6 - Low	54/145 = 0.37	1/25 = 0.040

The conditional probabilities in Table 11 were then adjusted to yield the "Test" A matrix in Table 12. The adjustment was carried out so that each conditional probability was greater than zero. It was assumed that the UXO found in category 6 was uncharacteristic of the binning performance because of remnant magnetic effects in using inert UXO in seeding the site. This is discussed in McDonald et al. (2001).

Table 12. "Test" Matrix A Based on Adjusted Badlands Bombing Range Results

Category	Probability Assigned to Category i Given that the Location is Truly Not UXO	Probability Assigned to Category i Given that the Location is Truly UXO
1 -High	0.04	0.720
2	0.07	0.200
3	0.25	0.038
4	0.02	0.001
5	0.25	0.040
6 - Low	0.37	0.001

It should be noted that the "Test" matrix A is not a random variable. Rather, it is the actual ability of the geophysical device and expert judgment to discriminate the locations into high and

low probability categories that contain UXO. The better the ability of the performance matrix to not place "true" UXO into the lower categories (false negatives), the "better" the discrimination method will perform and the sooner one will be able to stop digging. Thus, for purposes of illustration, the adjusted "Test" A performance matrix was used rather than the observed BBR performance matrix. If the observed performance matrix was used, the probability of UXO being present in the un-dug locations would seldom reach an acceptable level and 100% digging would usually be required because of the 0.04 probability in the low likelihood bin.

Stopping Rule

For this example, the stopping rule is determined by computing a lower 95% confidence bound on the probability that no UXO remains in the locations not in the initial dig set. If this 95% lower bound is greater than 95% after the data from the initial dig set is used, then digging up locations can cease. If the 95% lower bound computed from the initial dig set is less than 95%, then supplemental digging from the prioritized dig list is required. This supplemental digging scheme is based on answering the question: "If no UXO are found in the supplemental dig list, how many more dig locations are necessary until the 95% upper bound on the probability that no UXO remains in the un-dug locations is greater than 95%?" The supplemental dig list is drawn from the remaining un-dug locations in the "most likely UXO" category sequentially stepping down through the categories as needed. This is done so that in most situations the estimated lower bound will be as small as possible.

The stopping rule above is arbitrary and was chosen for demonstration purposes. The choice of whether to use a 90%, 95%, or 99% lower bound on the probability of no remaining UXO is a risk based decision, as is the choice of using a critical value of 5% to compare to the lower bound. This critical value could be any percentage that would provide an acceptable risk.

Example Results

Two examples are presented below using the performance matrix described as the "Test" A Matrix in Section 4.3.5 (Table 12). In each example a non-informative beta prior was used, i.e., $\alpha = \beta = 1$.

In the first example, assume that 490 anomaly locations were identified at a site through geophysical characterization. Also assume that the 490 locations are binned into the six bins as shown in Table 12. Assuming that roughly 5% of 490 locations were UXO, a simulated data set using the model of Section 4.3.2 yielded the true UXO/Non-UXO distribution shown in Table 13. In practice, this UXO/Non-UXO distribution would not be known until after digging but because this is simulated, we know which locations are UXO and which are not UXO.

Table 13. Binning and UXO Distribution in Simulated Site Assuming 5% UXO

Category	Total Locations	Count of Locations that are Truly Not UXO	Count of Locations that are Truly UXO
1 -High	37	19	18
2	38	33	5
3	117	116	1
4	10	10	0
5	116	116	0
6 - Low	172	172	0
Total	490	466	24

Several stages of sampling and analysis were considered to illustrate the methodology in Section 4.3.1. In the first stage all locations in bins 1-3 were dug up. In stages 2 – 4, locations were dug up in increments of 50 locations at a time, starting in bin 4, until the stopping rule was achieved. As illustrated in Table 14, the stopping rule was achieved in stage 4 after 342 locations were dug up (or about 70% of all locations).

Table 14. Results of Hypothetical Sampling Scheme on the 5% UXO Site

Stage	Number Dug Up	Total Dug Up	95% Confidence Lower Bound	Recommendation after Sampling Stage
1	(37, 38, 117, 0, 0, 0)	192 of 490 (39%)	0.04	Continue Digging
2	(37, 38, 117, 10, 40, 0)	242 of 490 (49%)	0.26	Continue Digging
3	(37, 38, 117, 10, 90, 0)	292 of 490 (60%)	0.67	Continue Digging
4	(37, 38, 117, 10, 116, 24)	342 of 490 (70%)	0.97	Stop Digging

In the second example again assume that 490 locations were identified at a site and that less than 1% of them were UXO. A simulated data set using the model of Section 4.3.2 yielded the site shown in Table 15.

Table 15. Binning and UXO Distribution in Simulated Site
Assuming 1% UXO

Category	Total Locations	Count of Locations that are Truly Not UXO	Count of Locations that are Truly UXO
1 -High	21	19	2
2	35	34	1
3	121	121	0
4	12	12	0
5	121	121	0
6 - Low	180	180	0
Total	490	487	3

Several stages of sampling and analysis were again considered to illustrate the methodology in Section 4.3.1. In the first stage all locations in bins 1-3 were dug up. In stages 2 – 3, locations were dug up in increments of 50 locations at a time, starting in bin 4. Because the 95% confidence lower bound was approaching the decision rule limit of 0.95 after the third stage, the last digging increment was 33 locations, the remaining locations in bin 5. In this case, as shown in Table 16, the stopping rule was achieved in stage 4 after 310 locations were dug up (or about 63% of all locations).

Table 16. Results of Hypothetical Sampling Scheme on the 1% UXO Site

Stage	Number Dug Up	Total Dug Up	95% Confidence Lower Bound	Recommendation After Sampling Stage
1	(21, 35, 121, 0, 0, 0)	177 of 490 (36%)	0.40	Continue Digging
2	(21, 35, 121, 12, 38, 0)	227 of 490 (46%)	0.76	Continue Digging
3	(21, 35, 121, 12, 88, 0)	277 of 490 (57%)	0.92	Continue Digging
4	(37, 38, 117, 12, 121, 0)	310 of 490 (63%)	0.99	Stop Digging

A slightly higher low bound was achieved in the second example in fewer stages with slightly fewer samples dug up. This is due to the fact that little UXO is present in the second six-bin example.

Further improvements should be achievable if the performance matrix had monotonically increasing and decreasing probabilities in columns 1 and 2 respectively. This would mean that fewer non-UXO locations would be binned in bins 1-3 and more would be binned in bins 4-6, thus eliminating needless sampling of locations that are non-UXO that were binned as more likely to be UXO. Addition gains, in terms of digging up fewer locations, could also be achieved in both examples if a stronger prior Beta distribution was used with $\alpha = 1$ and $\beta > 100$.

4.3.6 Summary and Conclusions

The number of locations that need to be dug up is dependent upon the binning performance matrix, the number of bins that are used for categorizing the anomalies, the sampling/digging scheme, and the stated stopping rule.

In the two examples presented in this paper, for the six category case, roughly a total of 60% to 70% of the locations needed to be dug up using the performance matrix presented and a stopping rule to insure that the 95% lower bound on the probability that no UXO remains in the un-dug locations is greater than 95%. The stopping rule used here was generic and for illustrative purposes only. If however, the conditions on the stopping rule are relaxed (i.e., lowering required confidence level or increasing the percent of UXO that can remain), then the expected number of digs required will decrease. On the other hand, if the stopping rule is made more stringent then the expected number of digs required will increase.

The Bayesian method used in the examples has the potential to significantly reduce the number of required digs if the prior Beta distribution is chosen so that the β parameter is 100 or greater, especially if an improved performance matrix were available. On the other hand, the Bayesian method requires some agreed upon prior distribution, which can affect the results. If the choice of the prior distribution is a difficult issue for decision makers to address, then the uninformed prior Beta distribution of $\alpha = \beta = 1$ used in this paper is the best choice because that distribution would yield results equivalent to a methodology that uses a maximum likelihood approach to estimate the probability of no UXO remaining in the un-dug locations.

The illustrated statistical procedure will work with any performance matrix, but the gains achieved over 100% sampling is most influenced by the performance matrix. Coupling this statistical methodology for stopping digging with improvements in UXO discrimination (reducing false negatives and positives) as well as improved geophysical calibration in the field will yield the biggest savings in terms of digging up unnecessary locations.

It is recommended that further research be conducted to evaluate the performance of the methodology and the sensitivity to the binning performance matrix, sequential digging schemes, number of bins, and confidence requirements. The methods should be programmed into easily used software (VSP) and tested in the field with recommendations for when the methods would be most/least beneficial for reducing unnecessary digging at a site.

4.4 Objective 4: Validation of Statistical Methods and Tools

Three simulated site demonstrations were performed. They are described in the following sections.

4.4.1 First Simulated Site Demonstration

A large Army firing range (10 km x 7 km) was simulated. It was assumed that the range had been used over a period of 50 years for battery targeting practice utilizing 155mm projectiles and 4.2-inch mortars. Potential target positions were expected to be randomly positioned within an impact area that is estimated to be located near the center of the range. An EM-61 Time Domain Metal Detector (TDMD) was selected as the instrument that would be used for conducting a survey of the range for locating potential UXO. In order to characterize the expected sensor performance, it was assumed that a preliminary baseline study was conducted with metal objects

designed to represent 155mm projectiles and 4.2 in. mortars buried at known depths. Results of the survey indicated a probability of detection (P_D) of 90% and a false alarm ratio (FAR) of 1.96. The site map is shown in Figure 16.

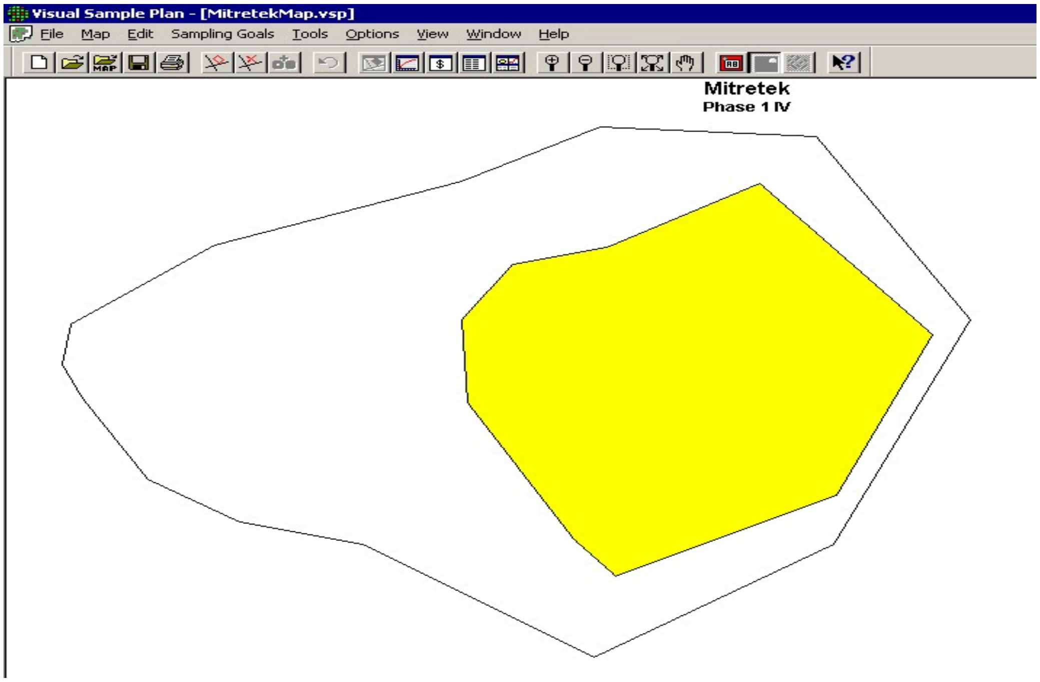

Figure 16. Map of the First Simulated Site Demonstration

To determine the transect spacing required, the following assumptions were derived from firing tables and lethal area dispersion calculations and used to determine the typical dispersion pattern for a 4.2" mortar. The probable errors in range and deflection from firing tables (FM 7-90, FM 23-91) are 27m and 6m respectively and the lethal area is 600 square meters represented by a 28m diameter circle. This resulted in the elliptical target area of concern of 218 m x 70 m.

The flagged TAs and boundaries are shown in Figure 17 and the true UXO locations are shown in Figure 18.

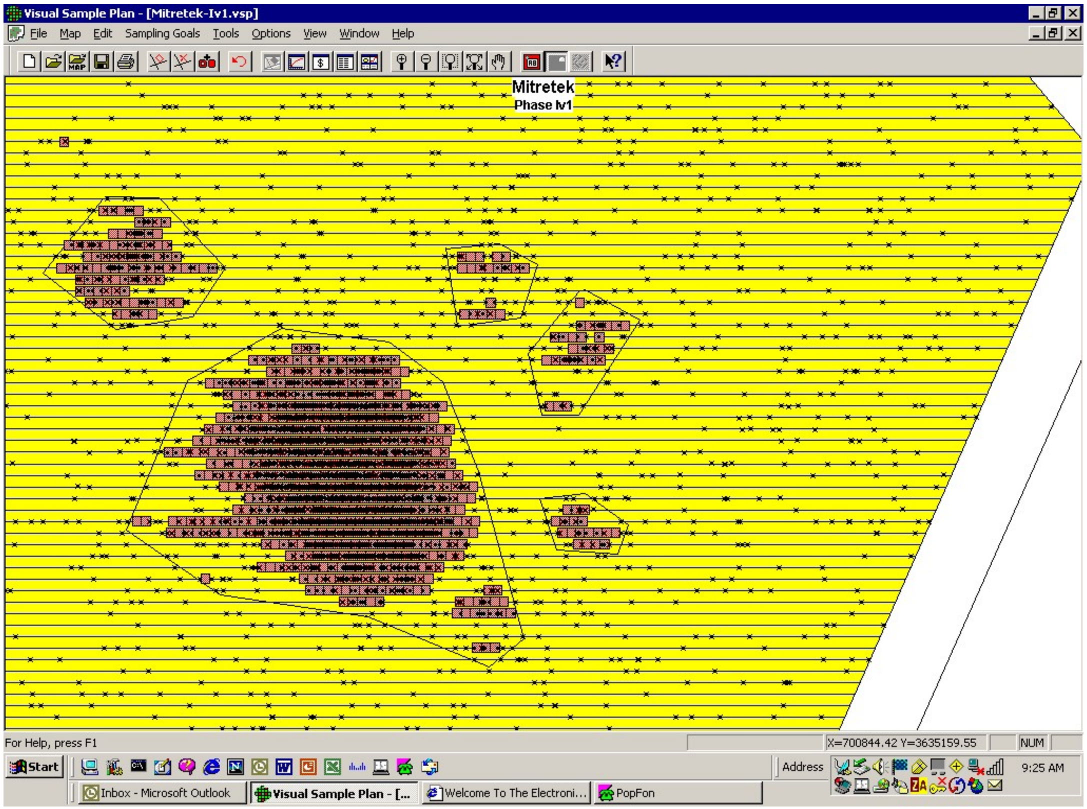

Figure 17. Flagged Target Areas and Boundaries for First Simulated Site Demonstration

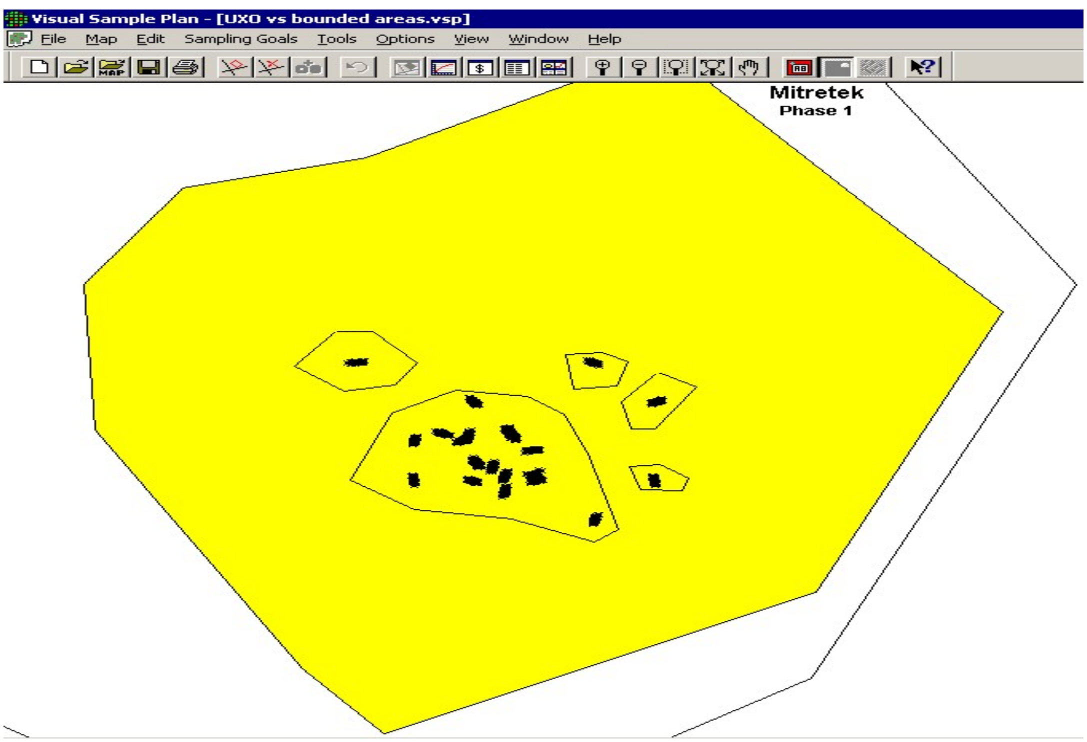

Figure 18. True UXO Locations for First Simulated Site Demonstration

The results from PNNL's portion of this demonstration show that

- 100% of targets were identified
- Derived target boundaries enclosed 100% of the UXO
- Isolated targets were found and delineated
- Assumed frag dispersion model may have been too conservative.

4.4.2 Second Simulated Site Demonstration

This site consisted of an Army firing range primarily used for battery targeting practice utilizing 155mm projectiles and 4.2 in. mortars over a 50 year period. A significant portion of the facility is believed to have been utilized as a training area and is thought to include a range with one or more impact areas located within the range.

The historical training area was estimated to cover an area of the installation approximately 21 km by 16 km. Historical information on the installation is minimal, however, there are indications that artillery and mortar firing points were located within the training area. Typically, munitions were fired into one or more impact areas located within the range. However their locations are unknown. In addition, no clear boundaries are known for the training area or the range. However, the archive search has provided some indication of the size and location of the training area. The site map is shown in Figure 19.

An EM-61 TDMD was selected as the instrument that would be used for conducting a survey of the training facility for locating potential UXO. In order to characterize the expected sensor performance, preliminary baseline studies were conducted in pre-selected 1 hectare grids with 10 metal objects designed to represent 155mm projectiles and 4.2 in. mortars buried at known depths in similar soil conditions as the suspect training area.

Results of the survey indicated a probability of detection (P_D) of between 60% and 90% and a false alarm ratio between 1.5 and 5.18. Many factors can affect the performance of the sensor. However, it is assumed for this exercise that these values represent the range of sensor performance across the terrain to be surveyed.

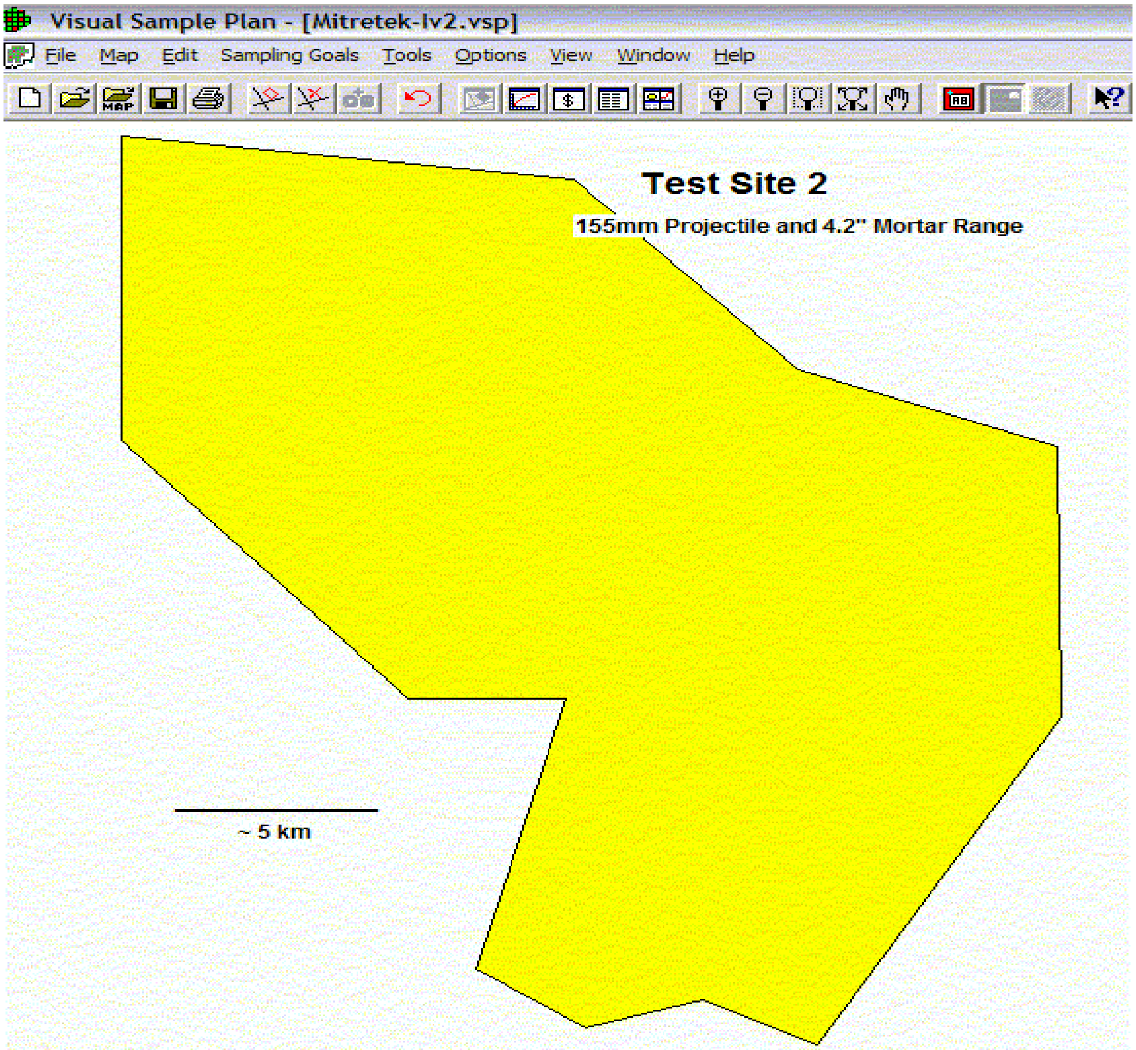

Figure 19. Map of Second Simulated Site Demonstration

To determine the transect spacing required, the following assumptions were derived from firing tables and used to determine the typical dispersion pattern for a 4.2 in. mortar. The same range and deflection parameters were used as those in the first simulated site demonstration, but this time the dispersion pattern was increased significantly (119m circles) to reflect calculated dispersion patterns for 50% of the frag materials, not just the lethal circles as shown in Figure 20.

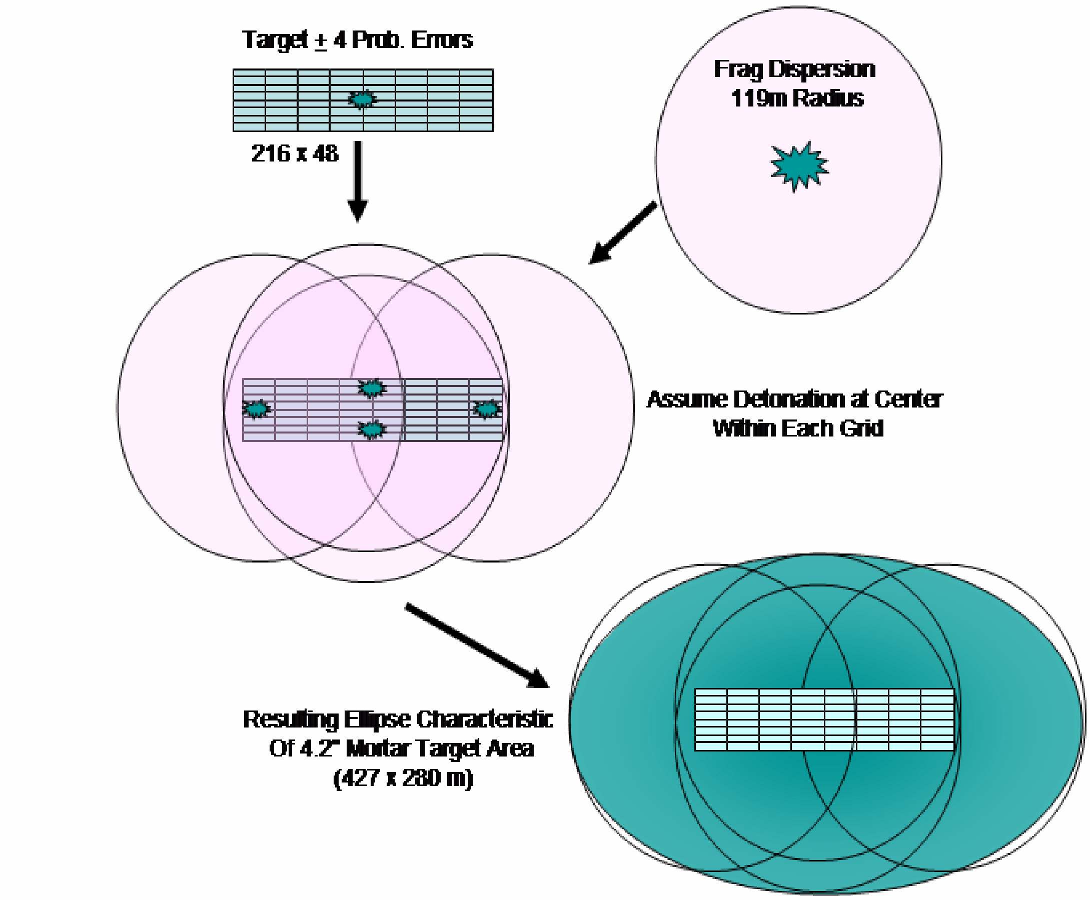

Figure 20. Diagram of Derivation of Target Area Size/Shape of Concern for the Second Simulated Site Demonstration

VSP was used to derive transect requirements. Then the transects were requested and suspected TAs were identified and delineated. Next, Mitretek, Inc. provided truth data. Figure 21 shows the true UXO locations (blue x's) encompassed by our bounded areas of concern.

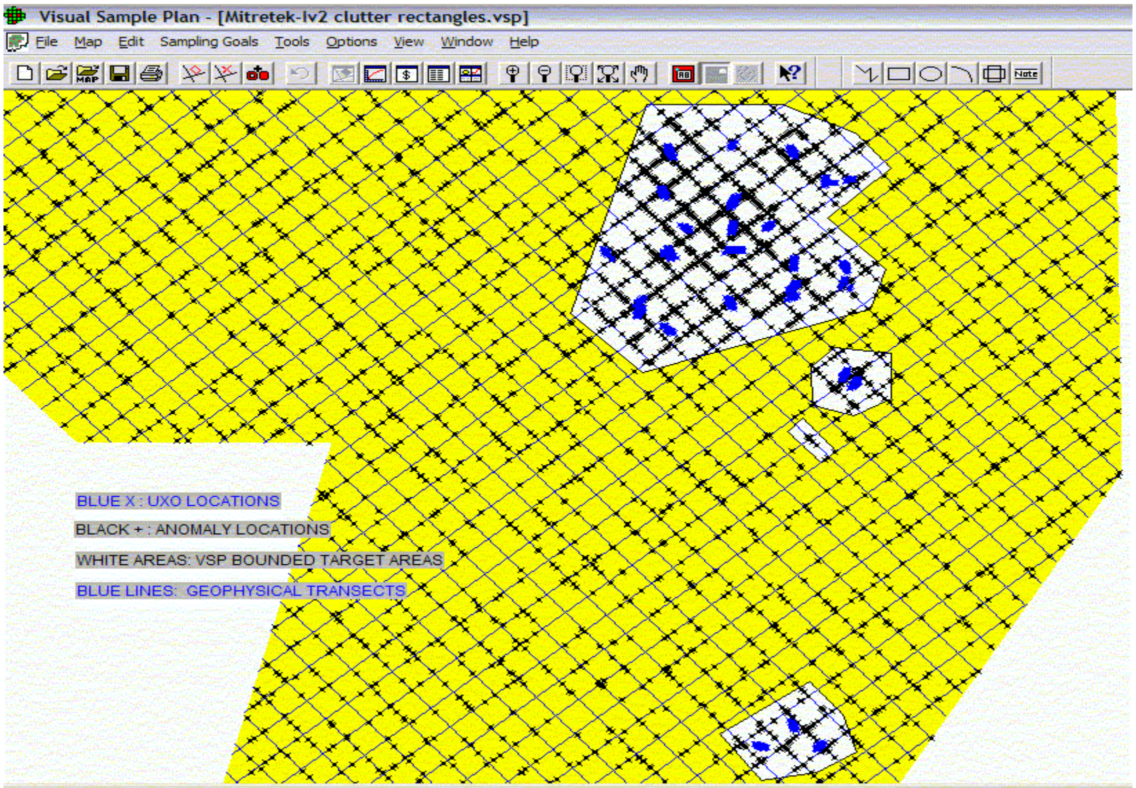

Figure 21. Map Showing Bounded Target Areas and UXO Clusters for the Second Simulated Site Demonstration

The results from PNNL's evaluation demonstrated that

- 100% of targets were identified
- Derived target boundaries enclosed 100% of the UXO
- Isolated targets were found and delineated
- Tools can handle gridded transects oriented as defined by the user.

4.4.3 Third Simulated Site Demonstration

The site consists of an Army firing range primarily used for battery targeting practice utilizing 155mm projectiles and 4.2 inch mortars. A significant portion of the facility is believed to have been utilized as a training area and is thought to include a range with one or more impact areas located within the range.

The historical training area is estimated to cover an area of the installation approximately 11 kilometers by 12 kilometers. Historical information on the installation is minimal, however, there are indications that artillery and mortar firing points were located within the training area. Typically, munitions were fired into the Impact areas located within the range. No clear boundaries are provided for the training area or the range, however, the archive search has provided some indication of the boundaries of the range area and those boundaries are provided. The simulation assumes that the range has been used primarily for artillery training with 155mm projectiles and 4.2 in. mortars. These are estimated to have been fired over a 50 year time period from firing points located within the range area at random targets located within the suspect

impact area(s). Specific locations of the firing points and targets are unknown. The exact size and number of Impact areas used for training are unknown. Available historical information indicates that artillery and mortar firing positions used for training exercises were primarily located towards the center of the impact area(s), however, this information has not been confirmed as accurate.

An EM-61 Time Domain Metal Detector (TDMD) was selected as the instrument that would be used for conducting a survey of the training facility for locating potential UXO. In order to characterize the expected sensor performance, preliminary baseline studies were conducted in pre-selected 1 hectare grids with metal objects designed to represent 155mm projectiles and 4.2 in. mortars buried at known depths in similar soil conditions as the suspect training area.

Results of the survey indicated a probability of detection (P_D) of between 60% and 90% and a false alarm ratio (FAR) between 1.5 and 5.18. Many factors can affect the performance of the sensor, however, it is assumed for this exercise that these values represent the range of sensor performance across the various terrain to be surveyed. The equipment utilized for performing a transect survey is limited to an overall individual swath width of either 1m or 2m. The cost of conducting a survey that is 1m wide is estimated to be $10 per meter. A 2m wide transect is $20 per meter.

The site map is shown in Figure 22 below.

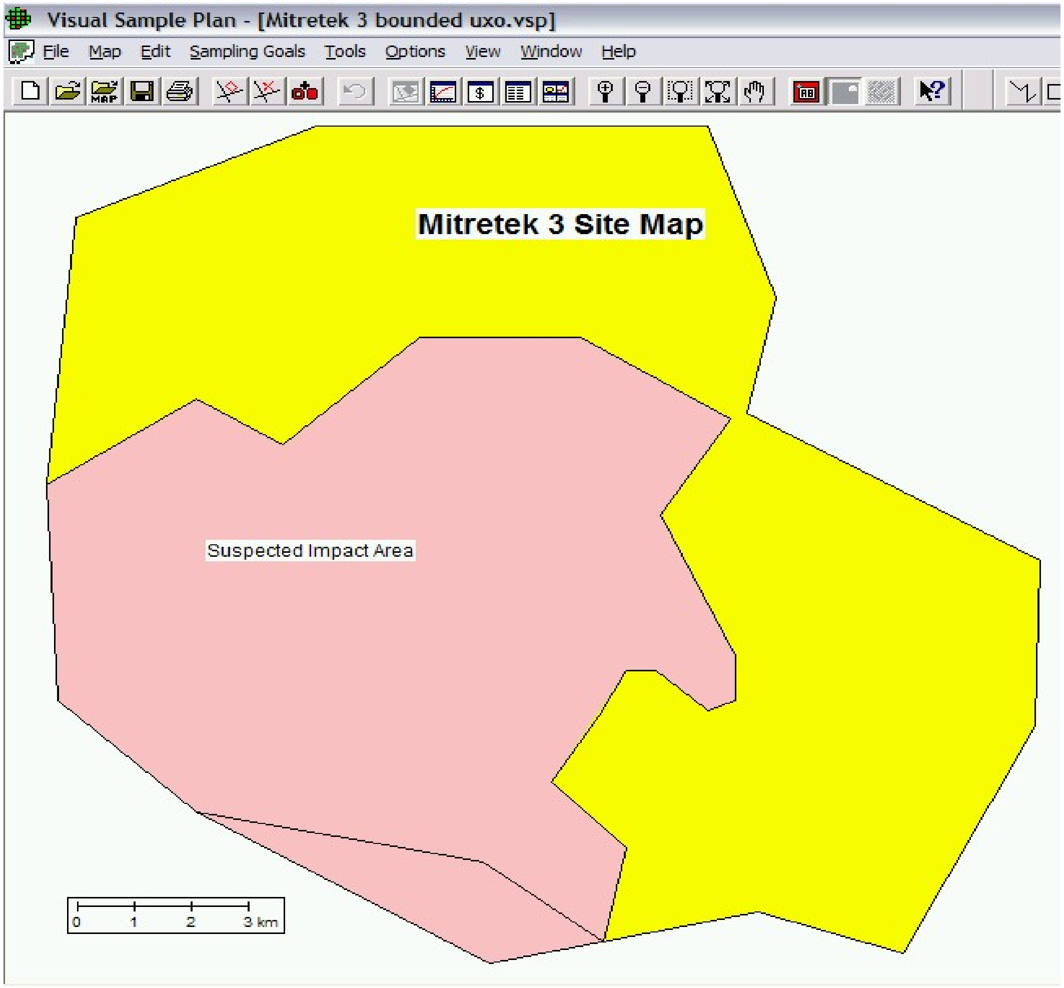

Figure 22. Map of the Third Simulated Site Demonstration

To determine the transect spacing required, the same assumptions were used as those shown for the second simulated site demonstration. We also developed a separate transect design for the suspected impact areas and the areas outside the suspected area. Inside the suspected area, we designed transects to ensure 98% probability of traversing and detecting TAs of concern if they existed. Outside the suspected area, we designed the transects to ensure 80% probability of traversing and detecting TAs of concern. After the anomaly data were received for the initial set of transects, additional transects were requested to better determine whether some potential TAs were indeed TAs of concern. The map with the original and added transects is shown in Figure 23. The resulting estimated target boundaries and actual UXO locations are shown in Figure 24.

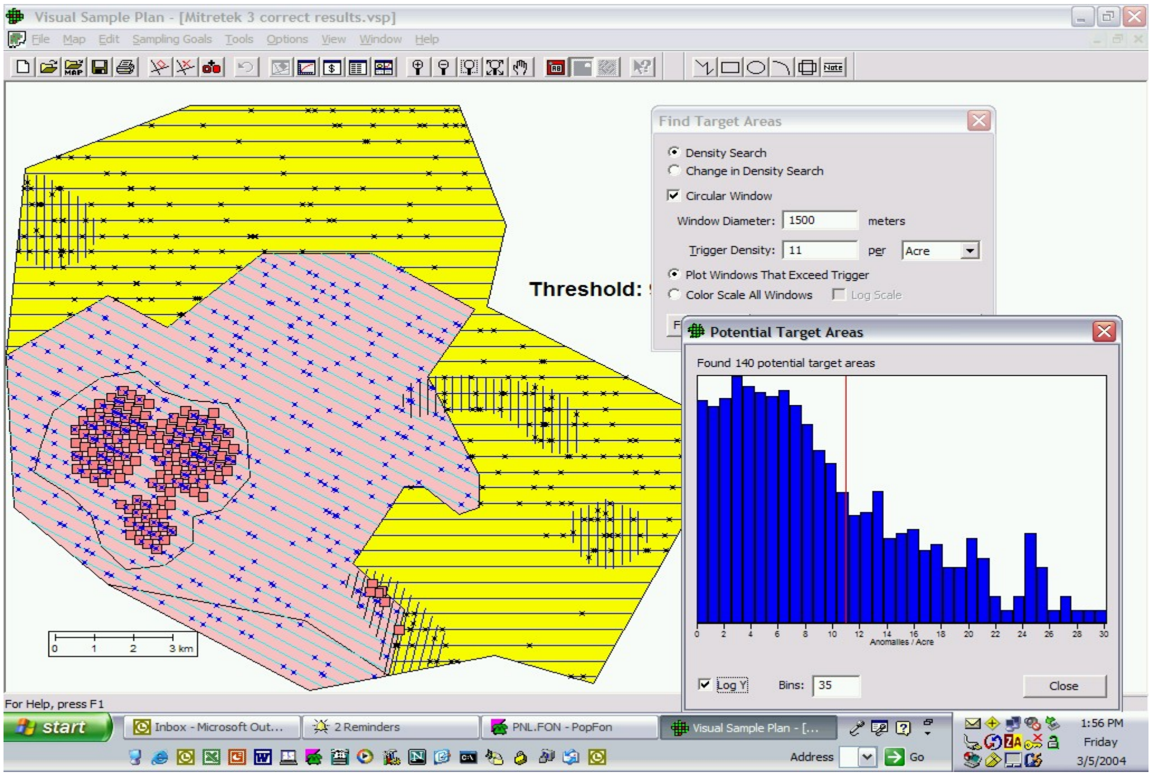

Figure 23: Anomalies Along Transects and Flagged Suspected Target Areas for the Third Simulated Site Demonstration

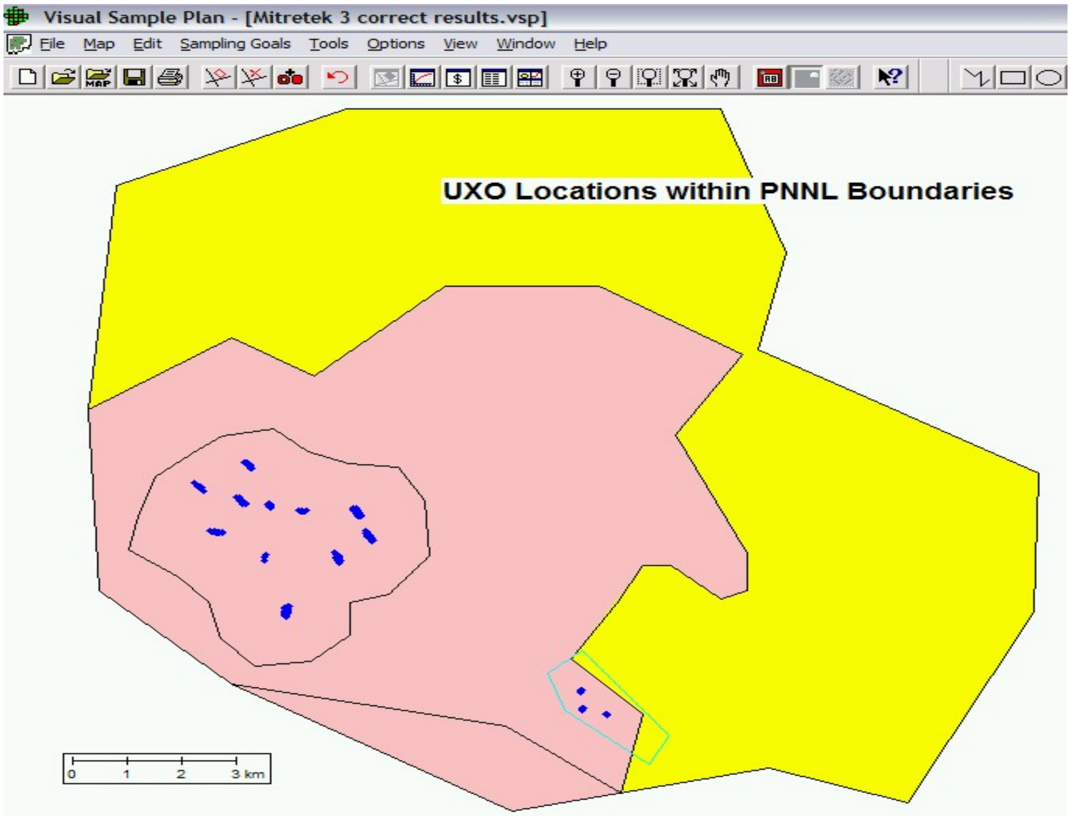

Figure 24: Bounded UXO Locations for the Third Simulated Site Demonstration

The results from PNNL's evaluation demonstrated that

- 100% of targets were identified
- Derived target boundaries enclosed 100% of the UXO
- Isolated targets were found and delineated
- Tools can handle multiple transect designs with different transect spacing and orientation.

5.0 Conclusions

This project has developed methods and tools to aid in finding unknown high-likelihood UXO areas, providing confidence that large un-surveyed areas are unlikely to contain UXO (i.e., are in compliance), and determining when it is acceptable to conclude that continued digging at locations on the dig list is unlikely to uncover any additional UXO. It is particularly significant that the target area detection and compliance methods developed have been incorporated into the Visual Sample Plan software code so that non-statisticians can implement the survey design methods in a systematic planning, DQO, framework. Also, the validation and demonstration efforts for the target area detection objectives conducted to date indicate that the survey design methods developed can indeed be used effectively to find target areas that are likely to contain UXO.

The close interaction of this PNNL project with the SNL SERDP project, which focused on the development of statistical anomaly density estimation methods, has produced an integrated methods and tools approach that is currently being demonstrated and evaluated with funding from ESTCP. Demonstration plans have been prepared for a validation exercise of the PNNL/SNL integrated approach on both a simulated site (using the SimRangE model developed by MTS) and also at the Isleta Pueblo S1 target area site in New Mexico. Demonstrations of the method at other actual training facilities are also anticipated.

The demonstrations of the methods are expected to result in improved design and analysis approaches. In addition, PNNL is currently reevaluating (for ESTCP) how the critical and trigger anomaly densities are defined and used in determining the required spacing of transects for detecting target areas. The idea is to explicitly define these density thresholds relative to the expected background clutter and to achieve a high probability of traversing, detecting, and flagging areas where the anomaly density is significantly higher than the average background density. This approach would be consistent with the DQO design process that is currently used in most other (non-UXO) modules in the VSP software.

Another need is to evaluate and demonstrate the utility and practicality of the Shilling and Wright/Grieve methods for determining the proportion of a site that must be surveyed in order to be confident that few if any UXO are present in large un-surveyed areas of a site or in areas that have undergone UXO removal actions. The Wright/Grieve approach is of particular interest because it involves the use of expert opinion to select a particular Beta distribution to model the degree of belief that an area does not contain UXO. Guidance on how this should be done should be developed.

It is important that those interested in using the methods receive training in the use of the VSP software. Although UXO modules in VSP are reasonably easy to use, a one or two-day course attended by practitioners is recommended. In addition to the UXO models in VSP, other current VSP modules might be used to design soil or sediment sampling studies to sample areas contaminated with explosives residues from detonation of various munition items. Indeed, one of the current VSP modules helps develop soil composite sampling designs to sample the perimeter of training areas to evaluate if munition residues in soil have migrated beyond the current boundary.

6.0 References

Casella, G., and R. L. Berger. 1990. *Statistical Inference*. Wadsworth. Beaumont, CA.

DoD. 1999. *Department of Defense Handbook, Companion Document to MIL-STD-1916*. MIL-HDBK-1916. U.S. Department of Defense, Washington, DC. February 1999.

Gilbert, R. O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Wiley & Sons, NY.

Gilbert, R. O., J. E. Wilson, R. F. O'Brien, D. K. Carlson, B. A. Pulsipher and D. J. Bates. 2003a. *Version 2.0 Visual Sample Plan (VSP): UXO Module Code Description and Verification*. PNNL-14267. Pacific Northwest National Laboratory, Richland, WA. April 2003.

Gilbert, R. O., J. E. Wilson, B. A. Pulsipher, D. K. Carlson, and L. Wrench. 2003b. *The Schilling and Wright/Grieve Methods for Assessing Compliance with UXO Requirements.* Pacific Northwest National Laboratory, Richland, WA. September 2003.

Grieve, A. P. 1994. "A Further Note on Sampling to Locate Rare Defectives with Strong Prior Evidence," *Biometrika* 81(4): 787-789.

Hassig, N. L., J. E. Wilson, R. O. Gilbert, D. K. Carlson, R. F. O'Brien, B. A. Pulsipher, C. A. McKinstry, and D. J. Bates. 2002. *Visual Sample Plan, Version 2.0 User's Guide*. PNNL-14002, Pacific Northwest National Laboratory, Richland, WA. September 2002.

Lantzer, N. and L. Wrench, K. J. Oates, H. Craig, T. Tomczyk, and S. Nelson. 2001. *Conceptual Site Model-Based Sampling Design*. The UXO/Countermine Forum 2001.

McDonald, J. R., D. Wright, N. Khadr, and H. H. Nelson. 2001. *Airborne NTADS Demonstration on the Impact Area of the Badlands Bombing Range.* NRL/PU/6110-02-453. Naval Research Laboratory, Washington DC.

O'Brien, R. F. 2002. *Probability of a Grid of Rectangular Transects of Positive Width Intersecting an Elliptical Target*, PNNL-13868. Pacific Northwest National Laboratory, Richland, WA.

O'Brien, R. F., P. G. Heasler, and K. K. Anderson. 2003. *A Statistical Bayesian Methodology to Stop Digging Geophysical Anomalies Based on Results from Incremental Digs*. Pacific Northwest National Laboratory, Richland, WA. August 2003.

Ostrouchov, G., G. P. Zimmerman, J. J. Beauchamp, V. V. Fedorov, and D. J. Downing. 1999. *Evaluation of Statistical Methodologies Used in U.S. Army Ordnance and Explosive Work."* ORNL/TM-13588. Oak Ridge National Laboratory, Oak Ridge, TN. September 1999.

Schilling, E. G. 1982. *Acceptance Sampling in Quality Control,* Marcel Dekker, New York

Simmons, M. 2003. *SimRangE Development and Analysis Plan (Final Draft)*, Mitretek Technical Report, Mitretek Systems, Inc. Falls Church, VA.

Singh, A., M. Engelhardt, and A. K. Singh. 2000. *Evaluation of U.S. Army Corps of Engineers Statistical UXO Sampling and Characterization Methodologies.* An internal EPA report, July 2000.

Singh, A., M. Engelhardt, and A. K. Singh. 2001. *Summary of Recent Results on SiteStats Evaluation Performed After the August 9-10, 2000, Partnership Meeting Between USACE and EPA NERL Las Vegas.* An internal EPA report.

USEPA. 2000. *Guidance for the Data Quality Objectives Process, EPA QA/G-4*. EPA/600/R-96/055. Office of Environmental Information, U.S. Environmental Protection Agency, Washington, DC. August 2000.

Wright, T. 1992. "A Note on Sampling to Locate Rare Defectives with Strong Prior Evidence," *Biometrika* 79(4):685-691.

7.0 Appendix

Technical Publications

<u>Technical Reports</u>

Bates, D. J. 2001. *Visual Sampling Plans and Unexploded Ordnance: Environmental Statistics in Action at PNNL.* PNNL-SA-34752. AMSTAT News (of the American Statistical Association): Vol. 288, pages 28-29, Pacific Northwest National Laboratory, Richland, WA. June 2001.

O'Brien, R. F. 2002. *Probabilities of a Grid of Rectangular Transects of Positive Width Intersecting an Elliptical Target*. PNNL-13868. Pacific Northwest National Laboratory, Richland, WA. April 2002.

Gilbert, R.O. 2003. *Version 2.0 Visual Sample Plan (VSP): UXO Module Code Description and Verification.* PNNL-14267. Pacific Northwest National Laboratory, Richland, WA. April 2003.

<u>Conference/Symposium Proceedings and/or Papers</u>

McKinstry, C. A., B. A. Pulsipher, and R. O. Gilbert. 2003. *Statistical Methods for Identifying and Bounding a UXO Target Area or Minefield.* PNNL-SA-39113. In International Conference on Requirements and Technologies for the Detection, Removal and Neutralization of Landmines and UXO. Pacific Northwest National Laboratory, Richland, WA. September 2003.

Pulsipher, B. A., R. O. Gilbert, and J. E. Wilson. 2002. *Statistical Tools for Designing Initial and Post-Removal UXO Characterization Surveys.* PNNL-SA-37405. Pacific Northwest National Laboratory, Richland, WA. In 9th UXO/Countermine Forum: Transformation in a Dynamic Environment, Orlando, Florida. September 2002

Pulsipher, B. A., J. E. Wilson, C. A. McKinstry, and R. O. Gilbert. 2003. *Identifying and Bounding a Minefield: Adapting Statistical Methods for UXO Target Detection.* PNNL-SA-40234. In Detection and Remediation Technologies for Mines and Minelike Targets VIII, Proceedings of the SPIE, Vol. 5089, pages 884-893. SPIE-Society for Optical Engineering, Bellingham, WA.

Pulsipher, B. A., J. E. Wilson, R. O. Gilbert, D. K. Carlson, R. F. O'Brien, G. A. Sandness, D. J. Bates, K. K. Anderson, and L. Wrench. 2004. *Finding and Delineating the Unknown Target Areas: A Statistical Tool for Designing UXO Transect Surveys*. Pacific Northwest National Laboratory, Richland, WA. UXO/Countermine Forum, St. Louis, Missouri. March 2004.

Gilbert, R. O., B. A. Pulsipher, J. E. Wilson, and L. Wrench. 2004. *Amount of Surveying Needed to Confidently Release DoD Sites*. PNNL-SA-40673. Pacific Northwest National Laboratory, Richland, WA. UXO/Countermine Forum, St. Louis, Missouri. March 2004.

PNNL White Papers

Gilbert, R. O., B. A. Pulsipher, D. K. Carlson, R. F. O'Brien, J. E. Wilson, D. J. Bates and G. A. Sandness. 2002. *Designing UXO Sensor Surveys and Using Visual Sample Plan (VSP) Software for Decision Making*. PNNL-SA-36792. Pacific Northwest National Laboratory, Richland, WA. March 2002.

Heasler,P. G., R. F. O'Brien, and K. K. Anderson. 2002. *Potential UXO Dig Stopping Methodology Based on Results from an Initial Sub-Sample*. Pacific Northwest National Laboratory, Richland, WA. July 2002.

Gilbert, R. O., J. E. Wilson, B. A. Pulsipher, D. K. Carlson and L. Wrench. 2003. *The Schilling and Wright/Grieve Methods for Assessing Compliance with UXO Requirements.* Pacific Northwest National Laboratory, Richland, WA. August 2003.

O'Brien, R. F., and P. G. Heasler. 2003. *Methodology to Stop Sampling Based on Results from an Initial Sub-Sample*. Pacific Northwest National Laboratory, Richland, WA. May 2003.

O'Brien, R. F., P. G. Heasler, and K. K. Anderson. 2003. *A Statistical Bayesian Methodology to Stop Digging Geophysical Anomalies Based on Results from Incremental Digs*. Pacific Northwest National Laboratory, Richland, WA. August 2003

Abstracts

Pulsipher, B. A. 2001. *Optimal Transect Spacing, Statistical Sampling, and Bayesian Compliance Methods.* POSTER. Pacific Northwest National Laboratory, Richland, WA. SERDP/ESTCP Partners in Environmental Technology Technical Symposium & Workshop, Washington, DC. November 2001.

Gilbert, R. O., B. A. Pulsipher, J. E. Wilson, D. K. Carlson, and R. F. O'Brien. 2001. *Statistical Methods and Tools for UXO Characterization*. POSTER. PNNL-SA-35637. Pacific Northwest National Laboratory, Richland, WA. SERDP/ESTCP Partners in Environmental Technology Technical Symposium and Workshop, Washington, DC. December 2001.

Pulsipher, B. A., R. O. Gilbert, J. E. Wilson, D. K. Carlson, R. F. O'Brien, G. A. Sandness, and D. J. Bates. 2002. *A DQO Visual Sample Plan Software Approach to Design of UXO Surveys*. PNNL-SA-37329, Pacific Northwest National Laboratory, Richland, WA. SERDP/ESTCP Partners in Environmental Technology Technical Symposium & Workshop, Washington, DC. December 2002.

Pulsipher, B. A., R. O. Gilbert, J. E. Wilson, R. F. O'Brien, D. J. Bates and G. A. Sandness. 2002. *Statistical Tools for Designing Initial and Post-Removal UXO Characterization Surveys*, POSTER, PNNL-SA-36796, Pacific Northwest National Laboratory, Richland, WA. SERDP/ESTCP Partners in Environmental Technology Technical Symposium & Workshop, Washington, DC. December 2002.

Pulsipher, B. A. 2003. *Visual Sample Plan (VSP): A Tool for Sampling/Survey Design and Residual Risk Evaluation*, PNNL-SA-38070. Pacific Northwest National Laboratory, Richland, WA. 5th Environmental Technology Symposium and Workshop, Charlotte, NC. March 2003.

Pulsipher, B. A., R. O. Gilbert, J. E. Wilson, D. K. Carlson, R. F. O'Brien, C. A. McKinstry, G. A. Sandness, and L. Wrench. *UXO Site Geophysical Survey Transect Design: A Demonstration Using a Simulated Site and Statistical Methodologies.* POSTER. Pacific Northwest National Laboratory, Richland, WA. SERDP/ESTCP Partners in Environmental Technology Technical Symposium & Workshop, Washington, DC. December 2003.

McKenna, S. A., B. A. Pulsipher, H. Saito, and R. O. Gilbert. 2003. *Detection and Prediction: Combining Tools for UXO Site Characterization*. PNNL-SA-40359. Sandia National Laboratories, Albuquerque, New Mexico and Pacific Northwest National Laboratory, Richland, WA. SERDP/ESTCP Partners in Environmental Technology Technical Symposium & Workshop, Washington, DC. December 2003.

Distribution

No. of Copies	**No. of Copies**